CONSERVATION GENETICS OF THE "DIABLITO" POISON FROG

A Thesis Submitted to the

College of Graduate and Postdoctoral Studies

in Partial Fulfilment of the Requirements

for the Degree of Master of Science

in the Department of Biology

University of Saskatchewan

Saskatoon

By

ANDRES GARCES-VASCONEZ

Permission to Use

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Biology,
University of Saskatchewan, 112 Science Pl.
Saskatoon, Saskatchewan (S7N 5E2)

ABSTRACT

Human activity is the main responsible for the biodiversity decline causing high rates of species extinction. The last few decades genetic factors have played important roles in determining population viability and extinction risk. Therefore, the discipline of conservation genetics emerged, taking advantage of molecular genetics tools to assess variation and study the fate of populations which are defined and identified by their genetic constituency. It is estimated that one-third of all amphibian species are threatened with extinction, and Neotropical species of frogs are no exception. The populations of the Diablito poison frog (*Oophaga sylvatica*) have abruptly declined in the last few decades. As a result, this species almost extinct in the Southern Ecuadorian part of its range. To develop conservation actions, one needs to identify conservation units, by choosing or prioritizing genetically distinct and reproductively independent populations for protection. In this study, I used microsatellite and mithocondrial DNA datasets, combined with high resolution pictures and bioclimatic data to asses of the existence of any genetic, morphological and/or ecological boundaries, and to reconstruct the recent demography of the populations. My study reveled two mitochondrial lineages with significant intraspecific divergence in aposematic coloration patterns. Moreover, based on both genetic and ecological exchangeability criteria, I propose that Diablito frogs in Ecuador are best characterized as two single evolutionarily significant units. However, a better sampling and further analysis should be considered to settle the appropriate (IUCN) conservation status and conservation units in this charismatic species.

ACKNOWLEDGMENTS

First, I would like to thank my supervisor Dr. Jose Andres for trusting on me for the development of this research project. This work would have not been possible without his guidance, feedback and invaluable support. Moreover, I would like to thank to Dr. Chris Todd and Dr. Carlos Carvalho for their advice as members of the MSc committee. Also, I greatly thank the Department of Biology of the University of Saskatchewan for the scholarship and the NSERC Discovery Grant to Dr. Jose Andres which supported this research.

In a special manner, I would like to thank Dr. Andres Posso-Terranova for his constant help and advice for the learning of the laboratory techniques, data analysis and for his friendship which has been also important during this years of thesis. I cannot forget to thank the other lab mates Martha, Prasob and Mitch for their help and camaraderie. I would also like to thank Gillian Murza, Halyna Heisler and Pamela Hind to for their administrative and technical support.

I also thank all my friends for their support and because they were always aware of my progress of this research. Specially, I want to thank to Alexandra Vargas for her immensely and priceless help during the field work. I have also to thank her for the unconditional support and help in the most difficult times during the development of this thesis. Lastly but the most important thank is for my family for their unconditional love and support which were fundamentals for the development of this research. I will be forever grateful to all.

TABLE OF CONTENTS

PERMISSION TO USE i
ABSTRACT ii
ACKNOWLEDGMENTS iii
LIST OF TABLES vi
LIST OF FIGURES vii
LIST OF ABBREVIATIONS ix
CHAPTER 1 INTRODUCTION 1
1.1 Conservation biology: A genetic approach 1
1.2 Conservation in amphibians 4
1.3 Life History of the Diablito poison frog, *O. sylvatica* 6
1.4 The "Diablito" poison dart frog, *Oophaga sylvatica:* Taxonomic and conservation status 6
CHAPTER 2 MATERIALS AND METHODS 9
2.1 Specimen sampling and DNA extraction 9
2.2 Phenotypic analyses 12
2.3 Mitochondrial genetic markers: Amplification and sequencing 14
2.4 Nuclear genetic markers: microsatellite genotyping 15
2.5 Phylogenetic analyses 15
2.6 Population structure analyses 17
2.7 Genetic diversity 18
2.8 Demographic history 20
2.8 Ecological analyses 21
CHAPTER 3 RESULTS 23
3.1 Phenotypic variation 23
3.2 Phylogenetic patterns of COI variation 27
3.3 Population differentation 32
3.4 Genetic diversity: microsatellite data 34
3.5 Genetic diversity: mitochondrial data 38
3.6 Demographic history 39
3.6.1 Bottleneck 41
3.7 Environmental clustering and correlation among molecular, phenotypic and environmental variables 42

CHAPTER 4 DISCUSSION 45
4.1 Phenotypic differentation 45
4.2 Genetic differentation 46
4.3 Effective population size and population size change 48
4.4 Conservation and management 49
4.5 Conclusions and future directions 51
LIST OF REFERENCES 52
APPENDIX I. SUPPLEMENTARY TABLES AND FIGURES 58

LIST OF TABLES

Table 1. Information for each studied population: locality name, code, number of samples collected, geographic coordinates and altitude, number of individuals sequenced and genotyped. 11

Table 2. Description of the 26 morphological variables measured used to estimate morphological clusters..... 13

Table 3. Estimates of Evolutionary Divergence over mtDNA Sequence Pair between *Oophaga* groups. HKY distances are shown under the diagonal. Standard error estimates are above the diagonal..... 28

Table 4. Results from STRUCTURAMA analysis at a range of priors for shape and scale in *O. sylvatica* based in microsatellite dataset. The optimal number of inferred populations is shown in light grey..... 34

Table 5. Summary of nuclear and mitochondrial genetic diversity. 36

Table 6. AMOVA analysis results. The molecular variation is mostly retained within populations.…….....37

Table 7. Summary of bottleneck analysis based on microsatellite data. Significant values ($p<0.05$) of excess of heterozygosity per population is shown in bold……..41

Table 8. The nineteen bioclimatic variables divided in temperature and precipitation. . 43

Table 9. Mantel and partial mantel test between genetic, environmental and geographic variables… 43

LIST OF FIGURES

Figure 1. Map of the sampled localities of *O. sylvatica* in Ecuador and their geographical distribution. Individuals were found in lowlands and foothill rainforest (47 to 738m above sea level) in northwestern Ecuador. Frogs pictures show the phenotypic variation among geographical localities……10

Figure 2. a) Map of sampling localities and phenotypic variation among populations. b) Discriminant analysis of 26 phenotypical variables. Graphic shows the optimal number of k (k=5). Scatter plot of morphological clusters. Dots represent each individual and ellipses and colours to each morphological cluster. c) Representative examples of each morph 25

Figure 3. a) PCA performed based on 19 color variables. The two first components retain mostly the variation (PC1= 67%, PC2=21%). Dots represents each individual. Five groups were revealed. 1) Red ellipse encloses individuals from Alto Tambo and Lita populations (light spots) 2) Khaki encloses Pto. Quito and Quingue (red, with or without spots) 3) Green encloses Palo Amarillo (red-dark orange) 4) Blue encloses Bilsa (orange) 5) Purple encloses Tundaloma and Mediania populations (bright orange). b) PCA of size variables. The first component represents the 99.9% of the variation. The size variables formed three groups. 1) The largest individuals (Alto Tambo, Lita and Palo Amarillo). 2) Medium size individuals (Bilsa). 3) The smallest individuals (Mediania, Tundaloma, Pto. Quito and Quingue)..... 26

Figure 4. Phylogenetic relationship of five species of *Oophaga* based on mitochondrial DNA (CO1). Sequences of *D. ventrimaculatus* was used as an outgroup. Tree was built under Bayesian and Maximum Likehood criteria. Numbers in nodes represent posterior probabilities and bootstrap (only values > 70% are shown). Population codes as in Table 1..... 29

Figure 5. Mitochondrial (COI) haplotype network of individuals of *O. sylvatica.* Sequences of *O. lehmanni, O. histrionica, O. pumilio* and *O. granulifera* are included as control species. *D. ventrimaculatus* is included as outgroup of the *Oophaga* clade. Circles indicate haplotypes, size is directly proportional to the number of individuals sharing that haplotype. Colours refer to the geographic location of the population that originate haplotype. Pie charts represent the percentage of each population sharing the same haplotype. Little circles in the haplotypes connections indicate the number of mutational steps. Two clades are clearly differentiated within *O. sylvatica* group..... 30

Figure 6. Neighbor-joining tree based in Nei's genetic distances (DS). North and south populations are clearly differencied in two lineages. Mediania population is situated between the two groups. 31

Figure 7. a) Population structure in *O. sylvatica* from microsatellite data. Bayesian clustering shows two (K=2) an three population clusters (K=3), being K=3 the optimal value of K according ΔK method. For K=3, blue color represents the north populations. Red and green the south populations. Some admixture is observable in Mediania population. For K=2 red color represents north

populations and green the south ones. Graphic on the right represents the optimal value of K estimated via ΔK method. b) DAPC scatterplot for microsatellite data. The first two components are showed. The first component separates mostly the north and south populations and the second component shows a well differentiated separation between Puerto Quito and Bilsa populations. DAPC eigenvalues axes contribution are showed in the bars graphic Population Codes as in Table 1. 33

Figure 8. Microsatellite private allelic frequency for both north and south mtDNA lineages. Frequency of private alleles was higher in the south clade. 37

Figure 9. Contribution of mtDNA allelic richness for each population in O. sylvatica to the total gene diversity. White bars indicate the intrapopulation diversity contribution (Crs). Grey bars show the contribution to the population differentiation (Crd). Black dots represent the total allelic richness contribution to the diversity (Crt). Palo Amarillo population has the higher contribution to the diversity. Mediania population mostly contributes to the divergence. There is not positive contribution to the total allelic richness.... 39

Figure 10. a) Observed distribution of pairwise differences of CO1 haplotypes between individuals. North and south lineages were analyzed independently. The observed probability show an unimodal shape in both cases (black line). The model is consistent with a recent spatial expansion. b) Extended Bayesian skyline plot for north and south lineages. Both cases show the effective population size has remained constant across time with a light increment in the last 0.2 MY (approx.). Blue and orange line represents the mean and median value of effective population size. Yellow and black lines indicate the upper and lower 95% posterior probability interval 40

Figure 11. a) Cluster of environmental variables based on Euclidean distances. Localities are grouped in three environmental clusters. First cluster: Quingue, Tundaloma and Mediania. Second cluster: Lita, Palo Amarillo and Alto Tambo. Third cluster: Pto. Quito and Bilsa. b) PCA of environmental variables. Dots represents each locality. Precipitation and altitude mostly retained the variation (PC1= 84%, PC2=12%) 44

LIST OF ABBREVIATIONS

AMOVA	Analysis of Molecular Variance
BI	Bayesian Inference
CITES	Convention on International Trade in Endagered Species
Crt	Allele Richness In Divergence
Crs	Allele Richness In Diversity
Crd	Total Allele Richness Contribution
DAPC	Discriminant Analysis Of Principal Components
DS	Nei's Standard Genetic Distances
EBSP	Extended Bayesian Skyline Plot
ESS	Effective Sample Sizes
ESU	Evolutionary Significant Unit
GPS	Global Positioning System
HWE	Hardy-Weinberg Equilibrium
LD	Linkage Disequilibrium
MCC	Maximum Clade Credibility
ML	Maximum Likehood
NEF	Nikkon Electronic Format
PCA	Principal Components Analysis
SAP	Shrimp Alkaline Phosphatase
SSD	Sum of Squared Deviations
TIFF	Tagged Image File Format
UCLN	Uncorrelated Log-Normal
π	Pi

CHAPTER 1

INTRODUCTION

1.1 Conservation biology: A genetics approach.

Biological diversity is rapidly declining as a direct and/or indirect consequence of human behavior. A large number of species are already extinct, and the populations of many others have been reduced to levels where they risk extinction (Lacy, 1997). Views on the relative importance of genetic factors in population viability and wildlife conservation have varied over the last few decades (Primmer, 2009; Lande, 1988), at times taking a back seat to the effects of demographic and environmental stochasticity (Allendorf et al. 2012). However, it is now widely recognized that genetic factors as play interactive and synergistic roles with ecological factors in determining population viability and extinction risk (Primmer, 2009).

Conservation genetics, a science theoretically rooted in population and quantitative genetics, takes advantage of molecular genetics tools to assess variation and study the fate of populations which are defined and identified by their genetic constituency. Because patterns of DNA variation contain a record of demography (Schneider & Excoffier, 1999; Primmer, 2009), the analysis of genetic polymorphisms can be used to make inferences about the history of populations, a critical piece of information to assess the potential genetic problems of small populations. For example, while long-term small population sizes will lead to low risk of inbreeding depression but high genetic loads, historically large populations close to mutation-drift equilibrium are likely to have a low genetic load but will be susceptible to inbreeding depression after a bottleneck (Ortego et al. 2007). Inbred populations may suffer higher risks of extinction because high levels of homozygosity negatively affect different fitness-related traits including survival (Liberg et al. 2005), reproductive success (Ortego et

al. 2007), behavior (Allendorf et al. 2012), parasite and disease resistance (Meffe et al. 2009; Frankham et al. 2002).

Genetic data are not a mere descriptor of variability or demographic process but an important management tool for populations or species. In 1995, following the recommendation of three different conservation workshops, conservation managers translocated eight female pumas from Texas (*Puma concolor stanleyana*) to the swamplands of Southern Florida in attempt to reverse the indications of inbreeding depression observed in the 20-30 surviving individuals of the Florida panther (*P.c. coryi*) (Maehr & Caddick, 1995). Fifteen years after this genetic restoration the population of Florida panther tripled, and showed a significant reduction in the incidence of phenotypic traits such as cardiac defects and low sperm quality traditionally associated with inbreeding depression (Johnson et al. 2010). Following the success of this enterprise, similar genetic restauration programs have been taking place in a wide range of threatened species (Hostetler, 2012).

Genetic data are also useful in setting conservation priorities. Conservation managers might decide to preserve the leopard (*Panthera pardus*), but which of the 9 currently recognized subspecies (UICN, 2016) of leopard should be preserved? Which of the 27 distinct phenotypically distinct populations should be prioritize? Clearly, existing taxonomic groups do not cover a substantial portion of the biological diversity. Consequently, many agencies and governments have established policies and legislation to protect "intraspecific population segments (or units)" (United States Government, 1988). The question is, how to define these units? In a first attempt to answer this question Ryder (1986) defined Evolutionary Significant Units (ESUs) as populations that represent significant adaptive variation based on concordance between sets of data (genetic, ecological, etc.). This rather broad original definition has sparkled many more definitions over the last few decades. For example, while Moritz (1994) ignored adaptive genetic variation and defined ESUs are

populations that are reciprocally monophyletic for mtDNA haplotypes and that show significant divergence of allele frequencies at *neutral* nuclear loci, for Waples (1995), Dizon et al. (1992) and Crandall et al. (2000) genetic differentiation (*i.e.* isolation) *per se* is not enough and require evolutionary units to be ecologically and/or adaptively unique such that ESUs represent a reservoir of phenotypic and adaptive genetic and variation. As a result, all these latter definitions promote a two-step approach to define ESUs: (1) to test for isolation (restricted or no gene flow), and (2) to estimate adaptive divergence across multiple traits (e.g., morphological, behavioral, life history, geographic range, etc.). The longer the isolation and the more different the selection pressures, the more likely are populations to represent ESUs worthy of preservation.

A potential problem with this conservation approach is that by conserving ESUs we are only preserving the products of evolution rather the evolutionary processes that contribute to biological diversity. Biodiversity is not fixed in time and space and our ultimate (and possibly ideal) goal should be to conserve the factors responsible for evolution itself (genetic polymorphism, connections between populations, etc.). Instead of just finding ESUs, a better strategy might be to identify historical levels of gene flow among populations and design networks of populations that will preserve them. Currently, next-generation sequencing projects are being conducted in a number of non-model organisms including many threatened species (Zhang et al. 2011). Genome-wide single nucleotide polymorphism (SNP) allele frequency data can be used for demographic inference under models of that consider changes in population size (Allendorf et al. 2012). These methods can also be applied to more than one population and complex demographic scenarios using multidimensional frequency spectrums and Approximate Bayesian Computation (ABC) methods (Corander & Arjas, 2008). There is no doubt that these type of genomics approaches will open up a new range of possibilities in conservation biology.

1.2 Conservation in amphibians

Living amphibians (Class Amphibia, Subclass Lissamphibia) include frogs and toads (Order Anura, ~ 6700 species), newts and salamanders (Order Caudata, ~695 species), and caecilians (Order Gymnophiona, ~205 species; AmphibiaWeb, 2017). Despite the relatively small number of amphibian species, according to the International Union for Conservation of Nature (IUCN) Red List of Threatened Species (IUCN, 2016) there are nearly as many threatened amphibians as threatened birds and mammals put together (2030 species) (IUCN, 2016). As a result, different authors have estimated that between 30 to 40% of amphibians are in danger of extinction (Baillie, 2004; Wake & Vredenburg, 2008). The dramatic world-wide decline of amphibian populations has been documented the last few years. Numerous potential factors including habitat destruction and modification, diseases, pollution, and increase UV-radiation are likely to contribute to population declines (Allendorf et al. 2012). Amongst them, the infection disease caused by the aquatic fungal pathogen *Batrachochytrium dendrobatidis* (Daszak et al. 2004) is likely to be the most important single contributing factor. Driven its dispersal capacity (Lips et al. 2008; Duellman & Trueb, 1986) has been detected in at least 48% of the amphibian species studied worldwide (Olson et al. 2013). In light of this global decline it is critically important to assess the conservation status of amphibian populations particularly in Latin America and the Caribbean where the largest numbers of endangered amphibian species occur (Stuart et al. 2004). To date, relatively few studies have evaluated the relative impact of the populations declining in these regions and these studies have generally been limited in scope, either because they have evaluated a small number of species or focused in a restricted geographic region. Poison dart frogs (*Dendrobatide*), an endemic group native to the tropical rainforests of Central and South America, exemplify the current limitations of our knowledge. Of the 157 species of dendrobatids only a few species have been studied, most of them showing taxonomic

uncertainties which makes difficult to assess their conservation status (Santos et al. 2009; Posso-Terranova & Andres, 2016a). While the majority of dendrobatids rely on crypsis to avoid predators, some members of this family are aposematic (*i.e*. both brightly colored and chemically defended; Grant et al. 2006). Within this Family, the members of the genus *Oophaga* display an extraordinarily diversity of warning coloration which have make them important species models in ecology and evolution. Despite their importance in evolutionary research, the *Oophaga* populations are seeing rapid population declines tied to habitat loss, diseases and the illegal pet trade (reviewed by Posso-Terranova & Andres, 2016b).

The ultimate aim of this thesis is to contribute to the conservation of these species by characterizing the demography and diversity of one of less-known the South American species of this genus, the Diablito poison frog (*O. sylvatica*). Most of our knowledge of this group comes from the relatively well-studied Strawberry poison frog *O. pumilio* (reviewed by Roland et al. 2016). A series of detailed studies including analysis in phenotypic, phylogenetic and population structure studies have shown that *O. pumilio* presented two major clades, one corresponding to archipelago of Bocas del Toro (Panama) and the other to mainland populations (from Nicaragua to south-eastern border of Costa Rica). While the lineages inhabiting the islands of Bocas del Toro present extremely color polymorphism, the mainland populations show a more continuous and less variable coloration (reviewed by Gehara et al. 2013). Similarly, in South America, the harlequin poison frog (*Oophaga histrionica*) has long been suspected to be a complex of highly differentiated lineages. In fact, a recent study based on multivariate clustering of morphological, ecological and genetic data suggests that this complex can be split in five different species, three of them new to science (Posso-Terranova & Andres, 2016a). Overall, these studies suggest that the currently recognized species of this group encompass a diverse array of evolutionary lineages and operational conservation units.

1.3 Life History of the Diablito poison frog, *O. sylvatica*

The Diablito poison frog is a diurnal, terrestrial species, that inhabits the most Southern range of the Chocoan bioregion. Although it is mostly found in patches of relatively undisturbed (primary) forest, this species has been also able to survive in secondary forest or even crops, such as banana plantations. When inactive it takes refuge under litter or wooden logs (Ortiz et al. 2016). Males are territorial, but in comparison to other species of territorial poison frogs, their territories can be relatively close to each other (~ 5m from conspecific males). As in other members of this Genus, males express territoriality by showing a wide range of aggressive behaviors, including simultaneous calling, advancing and retreating, and wrestling (Silverstone, 1975; Summers et al. 1997). Their diet is mostly comprised of ants and mites (McGugan et al. 2016). Individual frogs, sequester alkaloid-derived chemicals from the eaten prey. As a result, the alkaloids found in the skin of individual frogs varies along a latitudinal gradient across populations (McGugan et al. 2016).

As all the other *Oophagas* studied thus far, this species shows one of the most sophisticated reproductive behaviors in the Dendrobatidae family. Mating occurs without amplexus. Instead, the male deposits its sperm before the female lays the eggs either on the ground or in phytotelmata (Ortiz et al. 2016). Fertilized eggs (4-20 eggs per clutch) are cared by the male and after hatchling the female carry individual tadpoles on their back to small pools of water in axils of bromeliads. Females returns around 15 days to deposit unfertilized eggs to feed tadpoles (Summers et al. 1997).

1.4 The "Diablito" poison dart frog, *Oophaga sylvatica:* Taxonomic and conservation status.

The Choco region considered as one of the most diverse terrestrial ecoregions in the planet, but has been highly threatened. In Ecuador, the herpetofauna in this forest is one of the richest in the Neotropics, only surpassed by the upper Amazon basin (Arteaga et. al

2016). Amongst them, the Diablito poison frog (*O. sylvatica*), a diurnal species which inhabits the lowland and submontane rainforest area of the Tumbes-Choco-Magdalena biodiversity hotspot (0 – 1000 meters above sea level) (Coloma et al. 2004). The distribution of this species comprises southwest Colombia (deparments of Cauca and Nariño) and northwest Ecuador, where it has been reported in the provinces of Esmeraldas, Pichincha, Imbabura, Cotopaxi, Manabi, Santo Domingo de los Tsachilas, and Los Rios.

The Diablito poison frog, first described as a geographic variant of *Dendrobates histrionicus* (currently *O. histrionica,* Silverstone, 1975) by (Funkhouser in 1956) and was later recognized as *Dendrobates sylvaticus* Lotters et al. (1999), and *O.sylvatica* (Grant et al. 2006), is considered as Near Threatened by the IUCN red list of threatened species because is believed to be widely distributed in Colombia where distinct lineages of *Oophaga* (possibly *sylvatica*) are known to occur in several protected areas. However, the known distribution of the taxonomically well characterized populations of Ecuador has apparently shrunk sharply, and now they are possibly only abundant in the most northern part of its range and just overlap two protected areas (Reserva Ecológica Cotacachi-Cayapas and Parque Nacional Mache-Chindul). Because of this, and because this species is readly sought-after by the (illegal) pet market, *O.sylvatica* has been considered of Concern (Coloma et al. 2004) and has been included in the Convention on International Trade in Endagered Species of wild fauna and flora (CITES) where it is categorized in the Appendix II (i.e. species which trade must be controlled). As many other species, the major identified threats are the deforestation for agricultural development, illegal crops, logging, human settlement and illegal pet trade (Coloma et al. 2004).

In this thesis I focus on the conservation genetics of *O. sylvatica,* a polymorphic species for which we lack reliable information about its demography, genetic structure and evolutionary history. Current species descriptions suggest a high degree of phenotypic

variation similar to that observed in other *Oophaga* species of broad geographic distributions. I used microsatellite and mithocondrial DNA datasets, combined with high resolution pictures and bioclimatic data to assess the existence of any genetic, morphological and/or ecological boundaries, and to reconstruct the recent demography of the populations. This, in turn, allowed me to determine the number of operational conservation units, a key parameter for the development of future conservation strategies. Accordingly my objectives are the following:

1. To unravel the phyleogenetic relationships and the genetic constitution of populations of *O. sylvatica* in Ecuador.

2. To assess the role of geography, ecology and the genetic background in phenotypic and genetic divergence

3. To determine the number of evolutionary significant units using a combined approach that includes molecular, phenotypic and ecological variables.

CHAPTER 2

MATERIALS AND METHODS

2.1 Specimen Sampling and DNA extraction

We sampled 159 individuals in 8 different geographic locations across the known geographic distribution of the *O. sylvatica* in Ecuador, in the biodiverse region known as Tumbes-Choco-Magdalena forest (Coloma, et. al 2004) (Figure 1; Table 1). Field work was conducted either in the morning (6:00 am to 12:00 pm) or in the late afternoon (4:00 pm until sunset). However, most of the specimens were collected early in the morning when the frogs were usually more active. For each population, we recorded geographic location data every 50m. using a handheld GPS unit. High-resolution images and tissue samples were then collected as follows: raw digital pictures (Nikkon Electronic Format: NEF) were taken using a Nikkon D60 digital camera against a Spectralon grey reflectance standard (Labsphere, Congleton, UK). Images were then linearized with respect to light intensity (Stevens et al. 2007) using Adobe Photoshop CS6 version 13.0, and stored as TIFF images (lossless compression) to preserve fidelity for further analysis. Tissue samples were obtained by toe-clipping (Phillott et al. 2011). Toe-clips were preserved in 95% ethanol until DNA extraction. Genomic DNA was extracted using the MasterPure™ DNA purification kit (Epicentre, Madison, WI, USA) following the manufacturer's recommendations. All samples were obtained under the collection permits granted to Andres Garces Vasconez by Pontificia Universidad Catolica del Ecuador Exportation permits for the DNA samples were obtained from Santiago Ron.

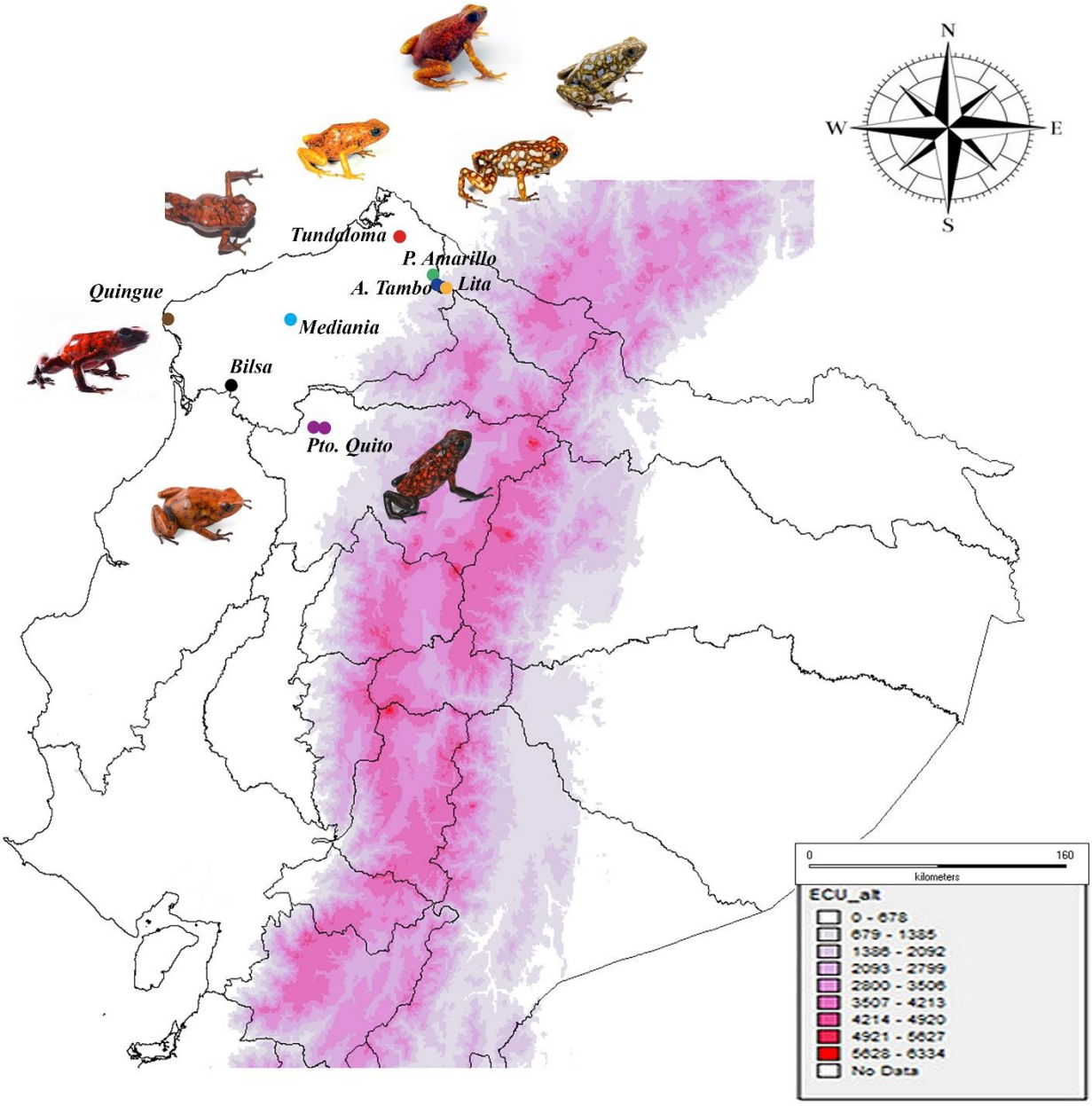

Figure 1. Map of the sampled localities of *O. sylvatica* in Ecuador and their geographical distribution. Individuals were found in lowlands and foothill rainforest (47 to 738m above sea level) in northwestern Ecuador. Frog pictures show the phenotypic variation among geographical localities.

Table 1. Information for each studied population: locality name, code, number of samples collected, geographic coordinates and altitude, number of individuals sequenced and genotyped

Locality/Population		Population Code	Number of Individuals	Coordinates		Altitude (m.o.s.l.)	COI (# sequences)	Microsatellites	
				Latitude (°N)	Longitude (°W)			# Loci	# Individuals
Lita	1	L	4	0.8993	-78.50776	547	6	13	10
	2		6	0.89364	-78.51616	615			
Tundaloma	1	TL	25	1.18278	-78.75253	47	7	13	31
	2		6	1.18287	-78.74966	122			
Palo Amarillo		PA	21	0.9601	-78.55795	663	3	13	21
Alto Tambo		AT	9	0.91494	-78.54166	738	8	13	9
Puerto Quito	1	PQ	22	0.10535	-79.18605	375	7	13	24
	2		3	0.11293	-79.23947	174			
Bilsa	1	B	12	0.34323	-79.71265	458	14	13	23
	2		11	0.34272	-79.71806	540			
Mediania		MD	20	0.71496	-79.37804	65	8	7	19
Quingue		Q	20	0.715447	-80.076366	169	6	7	18

2 Phenotypic analyses

I estimated the number of distinct morphological clusters using multivariate and discriminant grouping methods. Dorsal and ventral pictures of each specimen were scaled using ImageJ v. 1.48 (Sheffield, 2007) and a known metric reference (Posso-Terranova & Andres, 2016). For each specimen, a total of 26 variables (ndorsal=12; nventral=10; ndorsal/ventral=4) related to morphology (n= 7), pattern of coloration and coloration intensity (n= 19, Table 2) were evaluated. All variables were standardized prior to subsequent analyses. First, I performed a principal component analysis (PCA) as implemented in PAST v. 3.08 (Hammer et al. 2001). These resulting scores in the first two axes (~90% variation retained) were then used to perform a cluster analysis by applying the K-means algorithm implemented in the R-package mclust v 5. In addition, we also performed a discriminant analysis of principal components (DAPC) as implemented in the R package *Adegenet* v 2.0.0. (Jombart et al. 2010). This method has the advantage that allows for a probabilistic assignment of specimens to each group. To determine the optimal number of morphological clusters (K_M), I ran the sequential K-means (*find.clusters*, K=1-15) algorithm and identified one showing the lowest Bayesian Information Criterion (BIC) (Jombart et al. 2010).

Table 2. Description of the 26 morphological variables measured used to estimate morphological clusters

Position	Type	Variable	Description	Units
Dorsal/Ventral	Morphometric	Snout-cloaca length	Mean value of dorsal/ventral distance between the snout tip and cloaca	mm
Dorsal	Morphometric	Head width	Skull width between the ocular ridges	mm
Dorsal	Morphometric	Head depth (snout)	Depth of the head at its deepest point immediately posterior to the orbital cavities	mm
Dorsal/Ventral	Morphometric	Femur length	Dorsal/ventral mean value of femur length	mm
Dorsal/Ventral	Morphometric	Waist width	Dorsal/ventral mean value of waist width	mm
Dorsal/Ventral	Morphometric	Total area	Dorsal/ventral mean value of total area without including limbs	mm^2
Ventral	Morphometric	Jaw width	Distance between quadrate cartilages	mm
Dorsal	Coloration pattern	Number of colored dorsal spots	Dorsal count of spots with bright coloration, no including limbs	count
Ventral	Coloration pattern	Number of colored ventral spots	Ventral count of spots with bright coloration, no including limbs	count
Dorsal	Coloration pattern	Spot dorsal area	Area value of dorsal spots with bright coloration, no including limbs	mm^2
Ventral	Coloration pattern	Spot ventral area	Area value of ventral spots with bright coloration, no including limbs	mm^2
Dorsal	Coloration pattern	Background dorsal area	Area value of dorsal background, no including limbs	mm^2
Ventral	Coloration pattern	Background ventral area	Area value of ventral background, no including limbs	mm^2
Dorsal	Coloration pattern	Black hands	Categorical variable: presence/absence	No unit
Dorsal	Coloration intensity	Dorsal spot red intensity	R value (RGB component) dorsal spot coloration	RGB units
Dorsal	Coloration intensity	Dorsal spot green intensity	G value (RGB component) dorsal spot coloration	RGB units
Dorsal	Coloration intensity	Dorsal spot blue intensity	B value (RGB component) dorsal spot coloration	RGB units
Dorsal	Coloration intensity	Dorsal background red intensity	R value (RGB component) dorsal background coloration	RGB units
Dorsal	Coloration intensity	Dorsal background green intensity	G value (RGB component) dorsal background coloration	RGB units
Dorsal	Coloration intensity	Dorsal background blue intensity	B value (RGB component) dorsal background coloration	RGB units

Ventral	Coloration intensity	Ventral spot red intensity	R value (RGB component) ventral spot coloration	RGB units
Ventral	Coloration intensity	Ventral spot green intensity	G value (RGB component) ventral spot coloration	RGB units
Ventral	Coloration intensity	Ventral spot blue intensity	B value (RGB component) ventral spot coloration	RGB units
Ventral	Coloration intensity	Ventral background red intensity	R value (RGB component) ventral background coloration	RGB units
Ventral	Coloration intensity	Ventral background green intensity	G value (RGB component) ventral background coloration	RGB units
Ventral	Coloration intensity	Ventral background blue intensity	B value (RGB component) ventral background coloration	RGB units

2.3 Mitochondrial genetic markers: Amplification and sequencing

Polymerase chain reactions were performed with universal primer sequences COI-A (5'AGTATAAGC GTC TGG GTA GTC 3') and COI-F (5'CCT GCA GGA GGA GGA GAY CC 3') (Medina et. al 2013) to amplify 546 bp fragment of the mitochondrial gene Cytochrome c oxidase I (COI). Each PCR reaction (10µl) contained 15-25 ng of genomic DNA, 0.5 µM of each primer, 10 mM KCl, 2 mM $MgSO_4$ and 2.5 µl of PFU enzyme. The thermal profile consisted of an initial denaturation at 95°C for 5 min followed by 35 cycles of denaturation (94 °C for 50 s), annealing (55 °C for 45 s), elongation (72 °C for 1 min), and a final elongation at 72 °C for 5min. PCR products were purified by enzymatic treatment with 0.2 µl of Exonuclease I (10X), 0.2 µl of Shrimp Alkaline Phosphatase (SAP) (1,000 U/ml) and 1.6 µl Buffer 1X. Big Dye Terminator Cycle Sequencing Kits (Applied Biosystems) were used for cycle sequencing of the

purified amplicons, using the same primers as for PCR amplification (1 µl of purified amplicon, 0.5 µl 5X ABI Buffer, 0.7 µl RR Buffer, 0.15 µl 10 µM Primer, and 2.65 µl molecular grade water). Sequencing products from both directions were run on an ABI 31730XL DNA Analyzer with the ABI Data Collection Program.

2.4 Nuclear genetic markers: microsatellite genotyping

I selected thirteen microsatellite loci previously isolated for *O. histrionica* (*Ooph_109 A, Ooph_110 A, Ooph_44 A, Ooph_13 A, Ooph_6868 A, Ooph_63 A, A1, A2, A3, A4, A5, 3024* and *109.* (Posso-Terranova & Andres, 2016), which also cross-amplified in *O. sylvatica*. Initial screening of these loci and amplification optimization were done over a panel of ten individuals. Forward primers were fluorescent labelled and amplified for 157 individuals using polymerase chain reaction (PCR) with the following profile: one cycle of 240 s at 94 °C; five cycles at 94 °C, 56 °C, 72 °C of 30 s each step; five cycles of 94 °C, 54 °C, 72 °C of 30 s each step; thirty cycles of 30 s at 94 °C, 50 °C, 72 °C of 30 s each step; and final elongation for 5 min at 72 °C. PCR products were run in ABI-3130XL and genotyped GeneMapper Software 5.

2.5 Phylogenetic analyses

All mtDNA (COI) sequences were edited and aligned with DNASTAR Lasergene software: using SeqMan and MegAlign version 7.0.0. Edited sequences were then aligned using the Clustal-W method (Thompson et al. 1994). Aligned sequences (*Clustal-W)* were used to estimate the model of nucleotide substitution using the corrected Akaike Information Criterion as implemented in jModelTest v2.1.4 (Posada, 2008).

Gene genealogies were constructed using maximum likelihood (ML) and Bayesian inference (BI) as follows: first, I constructed BI-based genealogies using BEAST v. 2.1.2 (Drummond et al. 2012; Bouckaert et al. 2014) with a relaxed molecular clock and an uncorrelated log-normal (UCLN) model of molecular rate heterogeneity. I ran three chains for 10 million generations sampled every 100 steps. The resulting trees and log files of the three independent runs were combined using LOGCOMBINER 2.1.2. and resampled every 100 steps. For each estimated parameter, convergence was assessed using TRACER 1.6 (http://tree.bio.ed.ac.uk/software/tracer/) and effective samples sizes (ESS) were calculated to ensure adequate mixing (ESS>249, after 30% burn in). I summarized the posterior probability density of the combined tree and log files as a maximum clade credibility (MCC) tree using TREEANNOTATOR 2.1.2 (Bouckaert et al. 2014). I visualized all trees using FIGTREE 1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/). Second, I performed The ML tree searches using RAxML-HPC BlackBox Version 8 (Stamatakis, 2014). The bootstrapping was automatically halt by RAxML and the model to estimate the invariable sites was GTRGAMMA + I. Then, resulting trees were visualized and edited using the software FIGTREE 1.2.0 (http://tree.bio.ed.ac.uk/software/figtree/) and Photoshop CS6 version, respectively. Moreover, pairwise genetic distances (based on HKY model) were performed in Mega 7 software to determine the mtDNA genetic differentiation among *Oophaga* species. In addition, I built a haplotype network based on parsimony analysis, using the software TCS 1.21 (Clement et al. 2000). The analyzed sequences included *O. sylvatica* (n=59), *O. lehmanni* (n=3, Posso et al 2016a), *O. histrionica* (n= 4, Posso et al. 2016a), *O. pumilio* (n= 5, accession numbers: EF597190, EF597191, EF597192, EF597199, EF597200), *O. granulifera* (n=1, AF097505) and

Dendrobates ventrimaculatus (n=5) which was used as an outgroup (accession numbers: AF097502, AF412438, AF412437, AF412436, AF482824).

Finally, I used microsatellite to compute *Nei's* standard genetic mean distances (*Ds*, Nei, 1972), using the software Genepop. Ds distances matrix between populations was calculated for pairwise population comparison and used to construct an unrooted neighbor-joining tree (using the R-package *ape*).

2.6 Population structure analyses

Initially, I obtained statistical genetic descriptors such as, allele diversity, expected (H_E) and observed (H_O) heterozygosity estimates, and analyses of linkage disequilibrium and tests of Hardy–Weinberg equilibrium (HWE) using Arlequin 3.5.2.2 and Genepop software (Excoffier et al. 2005; Raymond & Rousset, 1995). Then, to study the population genetic structure of the existence of possible genetic boundaries within *O. sylvatica*, I first examined the population genetic structure using the Bayesian clustering methods implemented in STRUCTURE version 3.2.1 and STRUCTURAMA (Guillot et al. 2005). These methods use multi-locus genotypes for the inference of the number (*K*) of clusters (populations) that minimize deviations from Hardy–Weinberg. Given the recent history of the sampled population(s), I expected relatively low genetic differentiation, among populations. Therefore, I assumed a Dirichlet (i.e. correlated frequency) model to determine the number and composition of genetically differentiated subpopulations. A potential problem with this approach is that populations often possess complex genetic structures that are not always well described by explicit, hierarchical genetic models the sampling schemes and deviations from random mating not related to barriers to gene flow can have a strong influence on the detection and interpretation of genetic structure. Thus, I

also examined the genetic differentiation of *O. sylvatica* populations using a model free, multivariate approach. Unlike Bayesian clustering (Pritchard et al. 2000; Guillot et al. 2005; Corander et al. 2008), these methods do not rely on explicit population genetics models, and they are preferable when many loci are available and the structure is subtle (Jombart et al. 2010; Reeves & Richards 2009). Specifically, I used the Discriminant Analysis of Principal Components [DAPC] described by (Jombart et al. 2010)] as implemented in *Adegenet* (R.2.11.1; R Development Core Team, 2011). This method is designed to reveal groups of genetically related individuals directly from genetic polymorphism data, rather than on notions of existing structure that are reliant on the assumptions of HWE. STRUCTURE v 2.3.4 was run for a series of clusters (K_{NUC}) ranging from 1 (panmixia) to 8 (maximum number of localities sampled) using the following parameters: admixture ancestry model and correlated allele frequencies, 10000 burn-in, 100000 MCMC repetitions, and 10 iterations of each *K*. 10 independent runs with 1000000 MCMC iterations (burn-in=10000). The optimal value of K was estimated via ΔK method described by Evanno et al. 2005 as implemented in STRUCTURE HARVESTER (Earl, 2011). STRUCTURAMA (Huelsenbeck et al. 2007) was also used to infer K_{NUC}, assuming that the number of populations and the expected prior number of populations were random variables at a wide range of priors for shape and scale (see results). For this analysis, I used 100000 Markov chain Monte Carlo repetitions with a sample frequency of 100.

2.7 Genetic diversity

Linkage disequilibrium and deviation from Hardy-Weinberg equilibrium (HWE) was analyzed using the program Genepop (Raymond & Rousset, 1995). Each locus was tested using 1000 Markov chain steps and 1000 dememorization steps, and calculate linkage disequilibrium

(LD) between all pairs of loci using a likelihood-ratio test with 10,000 permutations. Levels of genetic diversity were determined by calculating the number, frequency of private alleles and the inbreeding coefficient using GenAlEx 6.5 (Peakall & Smouse, 2012). This software was also used to calculate the molecular variance (AMOVA). For this analysis, the assumptions were that each geographic locality corresponds to one population, and both mtDNA lineages corresponds to two different regions (North and South). In addition, heterozygosity index and the number of alleles at different loci were calculated with Arlequin 3.5.2.2 (Excoffier et al. 2005)

To assess mtDNA genetic diversity, I estimated haplotype and allele richness. Because the number of alleles depends on sample size, I use the rarefaction approach of Petit et al. (1998) as implemented in Contrib 1.4 (https://www6.bordeaux-aquitaine.inra.fr/biog eco_eng/Scientific-Production/Computer-software/Contrib-Permut/Contrib). This method uses rarefied samples in a decompositional analysis that determines the contribution (*C*) of each population to the total allelic richness (*Crt*) in a sample *n* of populations by comparing the allele richness including all populations to that after removing each population in turn. Individual population contributions to *Crt* are then decomposed into two components: The component of total allele richness attributable to within population richness (*Crs*) and the component attributable to among population differentiation (*Crd*). The various values of *C* can be negative if a particular population's contribution to any measure of richness or divergence is less than the overall average, or positive if its values are higher than the overall average (Petit et al. 1998; Zachary et al. 2008). Rarefaction size was set to be the same as the smallest population number. Additionally, I also estimated other mtDNA nucleotide (π) and haplotype diversity using the software DnaSp 5.10 (Librado & Rozas, 2009).

2.8 Demographic history

I used a suite of methods for investigating the demographic history of *O. sylvatica* using either mitochondrial data or microsatellite data. For the mitochondrial data, I first ran neutrality tests (Tajima's D and Fu's Fs) to detect either the signature of selection or the signature of recent population expansion (Fu & Li, 1993). These tests assume that population expansion will produce an excess of rare singleton mutations. While Tajima's D compares the average number of pairwise differences versus the number of segregating sites, Fs tests this relationship using the distribution of haplotypes (Fu & Li, 1993). Second, I executed a Mismatch Distribution analysis for each mtDNA lineage as implemented in Arlequin 3.5.2.2 (Excoffier et. al 2005). Populations at demographic equilibrium are characterized by multimodal mismatch distribution while unimodal distributions have been interpreted as being due to past demographic expansions (Slatkin & Hudson 1991; Rogers & Harpending, 1992). However, spatial expansions can also lead to unimodal mismatch distributions if neighboring subpopulations exchange enough migrants (Ray et al. 2003; Excoffier, 2005). Therefore, to further test for population expansion I constructed Extended Bayesian Skyline Plot (EBSP) using BEAST 1.8.4. For the mismatch analyses goodness of fit tests were performed to test the validity of the sudden expansion model, using a parametric bootstrap approach based on the sum of square deviations (SSD, Schneider & Excoffier, 1999) between the observed and expected mismatch distributions. The raggedness index which measures the smoothness of the mismatch distribution, and the expansion parameter (τ) was also calculated for each distribution. For the Bayesian Skyline analyses I used linear EBSP models, and random local clock, which reportedly performs better than strict and relaxed clocks for most situations using intraspecific data (Ho & Shapiro, 2011). I ran 10000000 Markov

Chain Monte Carlo (MCMC) chains, logging every 1000 chains. All other parameters were kept at default value. The results of these analyses were visualized using Tracer 1.6 (http://tree.bio.ed.ac.uk/software/tracer/). For the microsatellite data set I used BOTTLENECK v 1.2.02 (Cornuet & Luikart, 1996) to detect any signals of recent past bottlenecks. This analysis tests whether there is an excess of heterozygosity as compared with the heterozygosity expected from the observed number of alleles at each locus, assuming mutation–drift equilibrium. Bottlenecked populations are predicted to show an excess of heterozygosity, as the number of alleles is more severely affected than heterozygosity by a bottleneck in population size. Expected values were determined using 1000 iterations of a two-phase mutation model (Piry et al. 1999), with 95% single-step mutations and the variance of multiple-step changes set to 30%. The probability of significant heterozygosity excess was determined using Wilcoxon's signed-rank test.

2.9 Ecological analyses

Ecological data for each sampled locality were extracted from 19 high resolution (30 arc-seconds) bioclimatic variables (www.worldclim.org) plus altitude data collected for each geographic coordinate taken during the field work. These data were used to explore ecological differences among the sampled populations. To do so I used hierarchical and multivariate approaches. The software PRIMER v 7.0.10 (Clarke & Gorley, 2015) was used to obtain a matrix with Euclidean distances based in the bioclimatic data to perform a hierarchical cluster analysis, the parameter used was group average mode. In addition, a multivariate analysis (PCA) was run in PAST v. 3.08 to determine the ecological variables which contribute more to the divergence. Geographic distances among sampled localities were calculated using the web server

http://boulter.com/gps/distance/. Mantel's and partial Mantel tests were then applied to determine the association between pairwise Nei's genetic distances (Ds) (Nei, 1972), geographic and environmental distances. These analysis were performed using the function *mantel* and *partial.mantel* of the Vegan R-package (Oksanen et al. 2013).

CHAPTER 3

RESULTS

3.1 Phenotypic variation

Populations of *O. sylvatica* in Ecuador showed a high degree of phenotypic variation across the sampled localities (Figure 2a). To characterize and define such variation I used a discriminant multivariate approach to determine the number of different morphs, which in turn may contribute to identify the evolutionary conservation units. The unsupervised DAPC analysis resulted in the identification of five of phenotypic clusters (*i.e.* morphs K_M= 5), each of them including specimens from either one or two localities. In this latter case, the localities were not necessarily in close geographic proximity (Figure 2a, b). The first of these morphs, *Morph-1* (populations of Lita and Alto Tambo), is characterized by its relatively big size (Mean Total Area= 405.2 mm^2 / Mean Body Length= 35.5 mm - Mean Total Area= 420.4 mm^2 / Mean Body Length= 36.6 mm), presence of white – grey light spots (Mean RGB= 51.5 – 58.2) covering the dorsal and ventral body as well as the limbs, and a high dorsal background coloration ranging from light- to dark-orange and brown (Mean RGB= 37.5 - 28.3). The second morph (*Morph-2*), corresponds to the specimens collected in the Southern populations of Puerto Quito and Quingue. These frogs are relatively small (Mean Total Area= 245.4 mm^2 / Mean Body Length= 28.2 mm - Mean Total Area= 281 mm^2 / Mean Body Length= 31.5 mm), show a red spotted pattern with variable dorsal background coloration ranging from dark red (Mean RGB= 48.6 – 60.1) to almost black (Mean RGB= 27.1 – 29.2, Figure 2c). In contrast, the specimens classified as *Morph-3* (Palo Amarillo), were big in size (Mean Total Area= 398.9 mm^2 / Mean Body Length= 35.1 mm) and showed a variegated dorsal coloration, of orange, red and brown (Mean RGB= 50.5, Figure

2c) colors. *Morph-4*, (Bilsa) is the intermediate size group (Mean Total Area= 306.5 mm^2 / Mean Body Length= 33.1 mm) and shows a relatively homogenous dorsal coloration of relative similar orange tones (Mean RGB= 34.3, Figure 2c). Individuals included in *Morph-5* (Tundaloma and Mediania) is the smallest one of all (Mean Total Area= 241.8 mm^2 / Mean Body Length= 26.9 mm - Mean Total Area= 248.9 mm^2 / Mean Body Length= 29 mm). Coloration patterns similar to those found in *Morph-4*, but with brighter orange (Mean RGB= 48.1 – 52.9). A PCA of color variables alone (Figure 3a) rendered a similar partition of phenotypic variation: A spotted, red and black cluster (Puerto Quito and Quingue), a light-grey spotted cluster (Lita and Alto Tambo), The dull (Bilsa) to bright (Mediania and Tundaloma) orange cluster(s), and the cluster of profusely variegated specimens (Palo Amarillo). A PCA analysis based on all other morphological measurements (see Table 2) revealed that other that color and coloration patterns body size is the single variable that explains most of the phenotypic variation. Accordingly, a first cluster, grouped the large specimens of the northern populations (Palo Amarillo, Lita and Alto Tambo) that are geographically close to each other. A second cluster, included all individuals collected Bilsa population which are of intermediate size, and third group included all the small specimens from either the northern population of Tundaloma or the southern populations of Mediania, Puerto Quito and Quingue (Figure 3b).

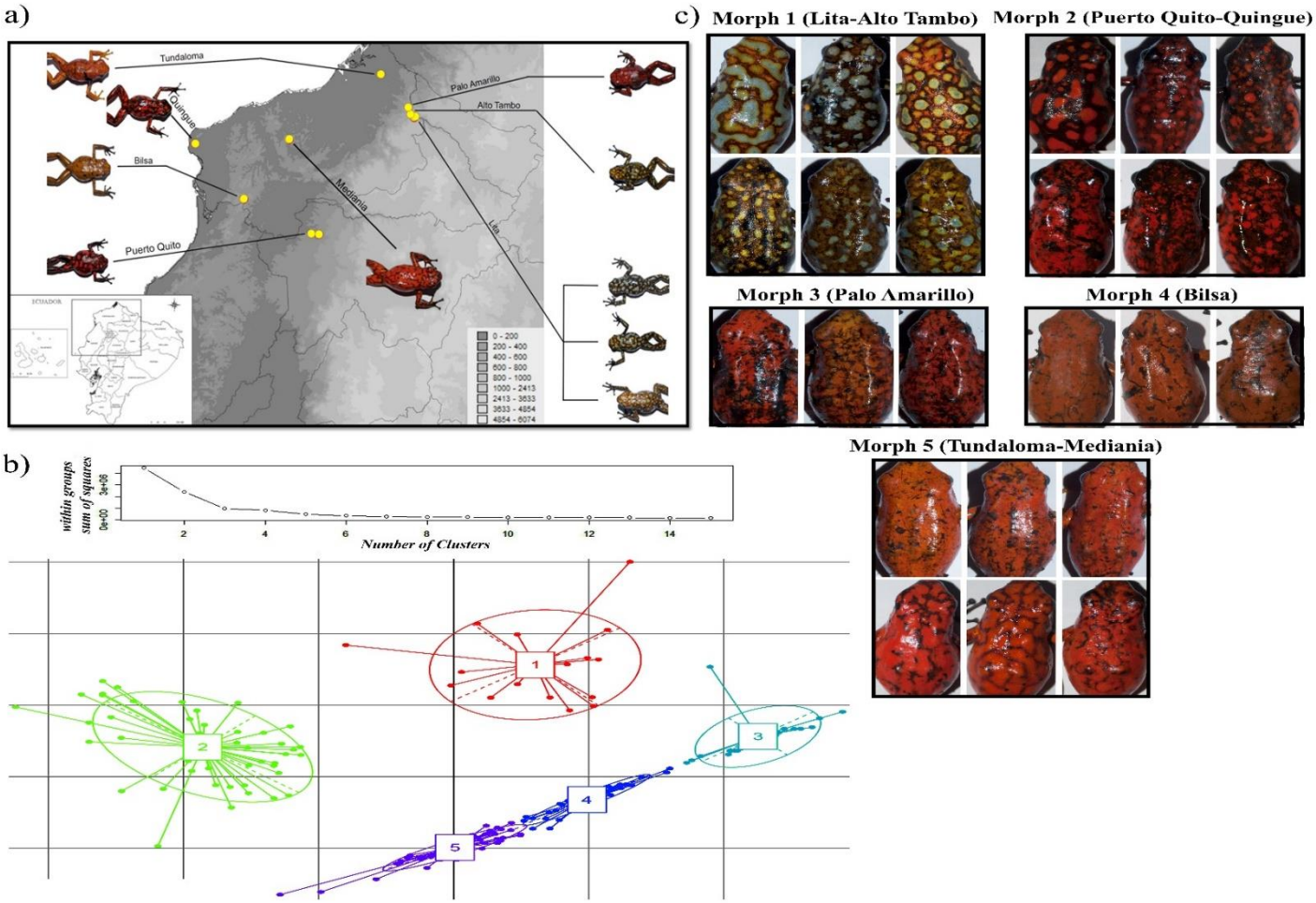

Figure 2. a) Map of sampling localities and phenotypic variation among populations. b) Discriminant analysis of 26 phenotypical variables. Graphic shows the optimal number of k (k=5). Scatter plot of morphological clusters. Dots represent each individual and ellipses and colours to each morphological cluster. c) Representative examples of each morph.

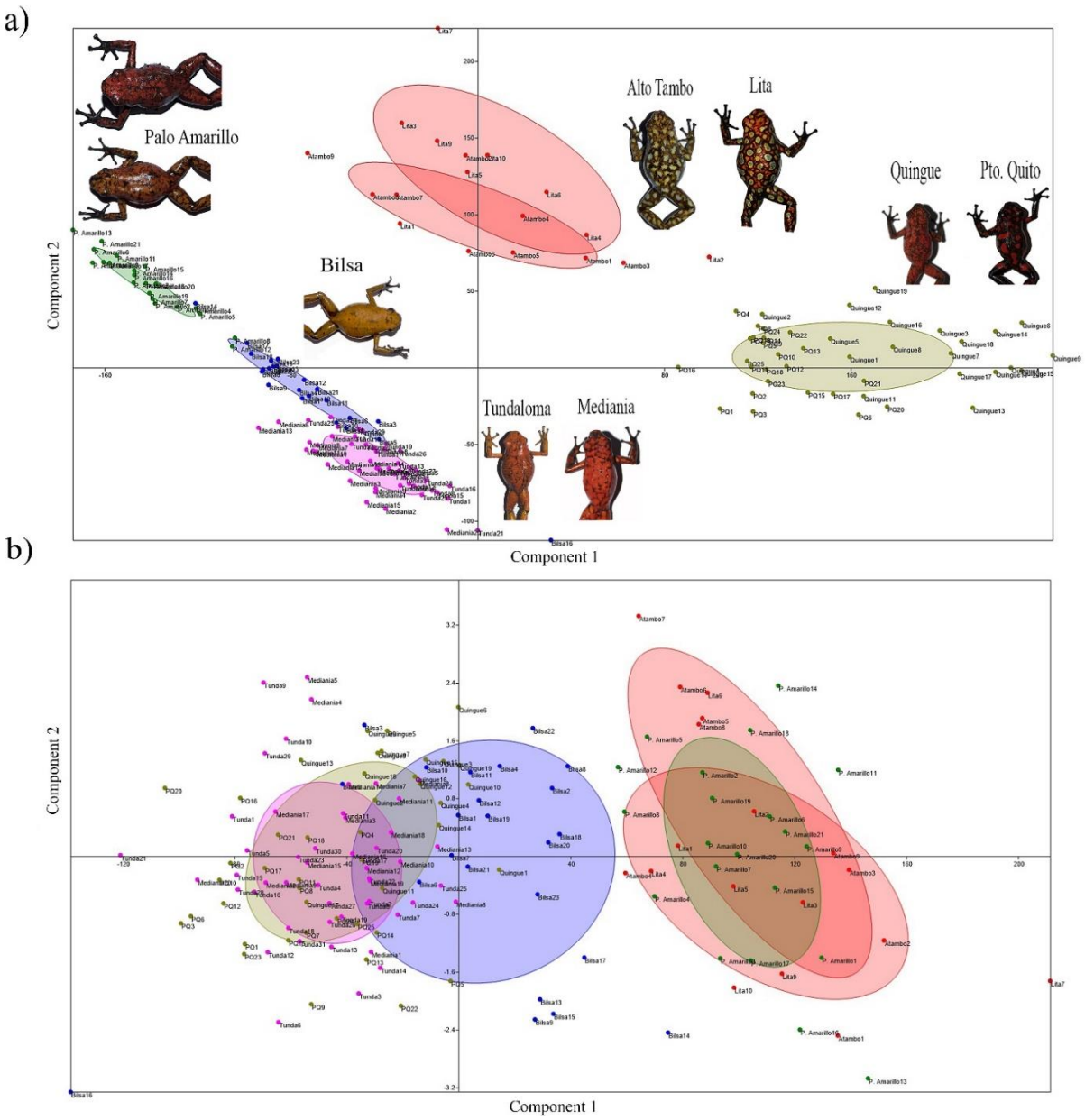

Figure 3. a) PCA performed based on 19 color variables. The two first components retain mostly the variation (PC1= 67%, PC2=21%). Dots represents each individual. Five groups were revealed. 1) Red ellipse encloses individuals from Alto Tambo and Lita populations (light spots) 2) Khaki encloses Pto. Quito and Quingue (red, with or without spots) 3) Green encloses Palo Amarillo (red-dark orange) 4) Blue encloses Bilsa (orange) 5) Purple encloses Tundaloma and Mediania populations (bright orange). b) PCA of size variables. The first component represents the 99.9% of the variation. The size variables formed three groups. 1) The largest individuals (Alto Tambo, Lita and Palo Amarillo). 2) Medium size individuals (Bilsa). 3) The smallest individuals (Mediania, Tundaloma, Pto. Quito and Quingue).

3.2 Phylogenetic patterns of COI variation

I analyzed 72 individuals from 5 different *Oophaga* species, representing 12 sites in the low-land tropical forest of the Chocoan biogeographic region. These sites include the eight populations of *O. sylvatica* sampled in Ecuador, one for *O. lehmanni* (La Cruceta) and *O. histrionica* (Santa Cecilia) in Colombia, as well as populations of Panama for *O. pumilio* (Bocas del Toro) and Costa Rica for *O. granulifera* (Corcovado). In *O. sylvatica,* I found 15 unique *COI* haplotypes, 546 bp in length, of which 22 (4%) nucleotide sites were polymorphic and 16 (2.9%) parsimony-informative. The ML and Bayesian phylogenies (Figure 4) had very similar topologies, both supporting the monophyly of the Ecuadorian populations of *O. sylvatica,* but dividing this species into two deep mitochondrial clades with strong bootstrap support (100%) and a high posterior probability. As expected, the South American *Oophaga* formed a monophyletic group sister to the Central American species (*O. pumilio* and *O. granulifera,* Figure 4). The (HKY model) pairwise genetic distances between nominal species fell into three groups: 1) rate distances between either *O. granulifera* and all other species clades (0.13 – 0.10); and 2) distances between *O. histrionica, O. sylvatica, O. pumilio* and *O. lehmanni* (0.04–0.07); 3) distances between *O. sylvatica* clades (0.02) (Table 3). Based upon the commonly used mitochondrial DNA (mtDNA) clock rate of ~2% pairwise divergence/Myr (Crawford, 2003), these two diverged around ~1.2 MY ago, during the Pleistocene. The two *O. sylvatica* clades showed geographic structure, one occurring in the northern Ecuador (Lita, Alto Tambo, Palo Amarillo and Tundaloma) and the other one in the in the South (Puerto Quito, Bilsa, Mediania and Quingue). Accordingly, the haplotype network showed the division of *O. sylvatica* in two haplogroups (North and South, Figure 5). Within the northern haplogroup, population of Lita

shows the highest level of haplotype diversity, including the presence of a single distinct mtDNA haplotype separated from all the rest by 11 mutational steps. Haplotype diversity is lower in the South. The star-like shape of this part of the network indicates very low levels of sequence divergence and a high frequency of unique mutations. Interestingly, all the individuals from Mediania (all included in the Southern haplogroup) carried the same private haplotype, not found in any of the other populations. An unrooted neighbour-joining tree based on nuclear markers (microsatellites and the standard genetic distance of Nei (Ds) (Nei, 1972), exhibited the same topology (Figure 6), where the localities from the North consistently cluster together with 100% bootstrap support, as do the localities from the South, the Mediania population falling between these two groupings.

Table 3. Estimates of Evolutionary Divergence over mtDNA Sequence Pair between *Oophaga* groups. HKY model pairwise distances are shown under the diagonal. Standard error estimates are above the diagonal.

	O. sylvatica (south clade)	*O. pumilio*	*O. sylvatica* (north clade)	*O. lehmanni*	*O. histrionica*	*O. granulifera*
O. sylvatica (south clade)		0.014	0.007	0.008	0.012	0.019
O. pumilio	0.085		0.014	0.012	0.012	0.016
O. sylvatica (north clade)	0.024	0.084		0.010	0.012	0.018
O. lehmanni	0.035	0.067	0.042		0.009	0.019
O. histrionica	0.062	0.071	0.063	0.044		0.020
O. granulifera	0.132	0.108	0.123	0.116	0.122	

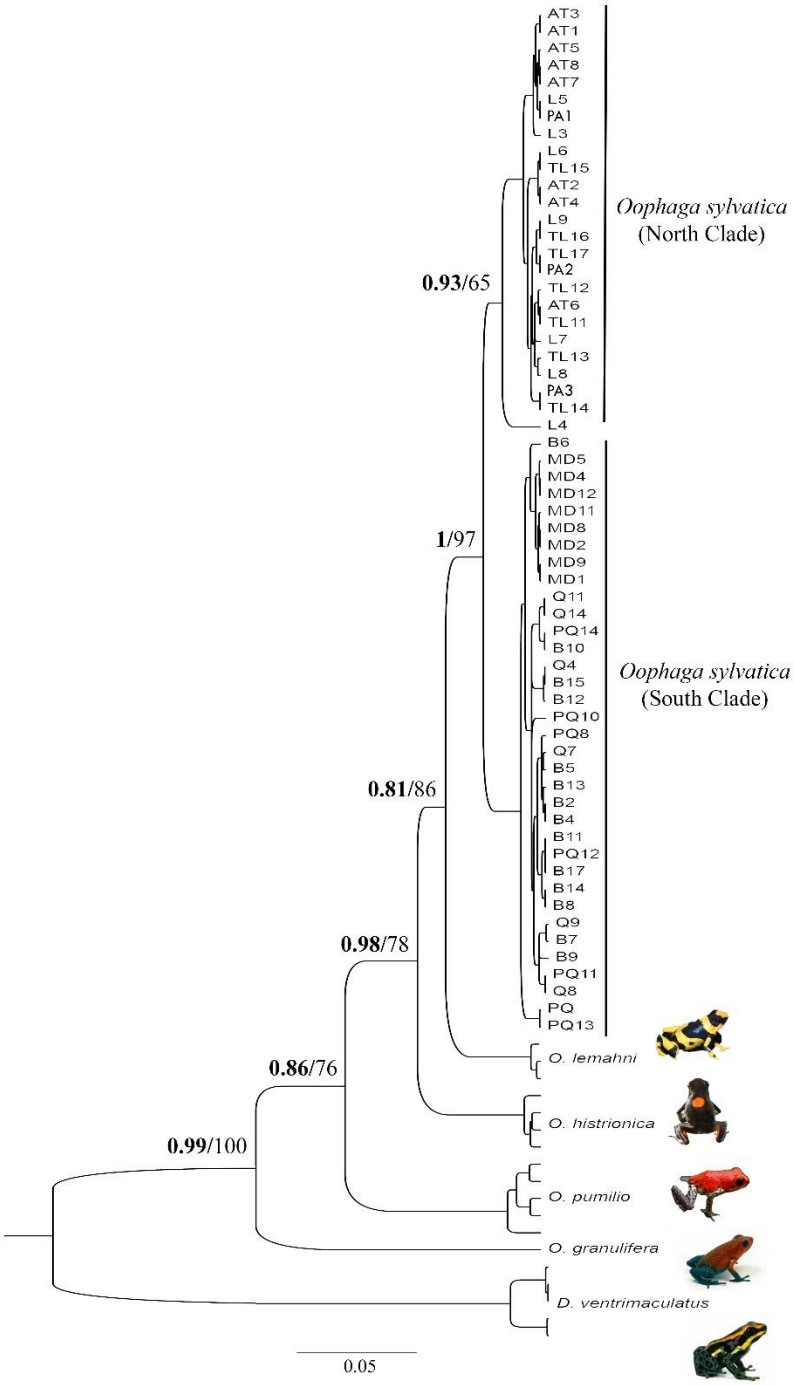

Figure 4. Phylogenetic relationship of five species of *Oophaga* based on mitochondrial DNA (CO1). Sequences of *D. ventrimaculatus* was used as an outgroup. Tree was built under Bayesian and Maximum Likehood criteria. Numbers in nodes represent posterior probabilities and bootstrap (only values > 70% are shown). Population codes as in Table 1.

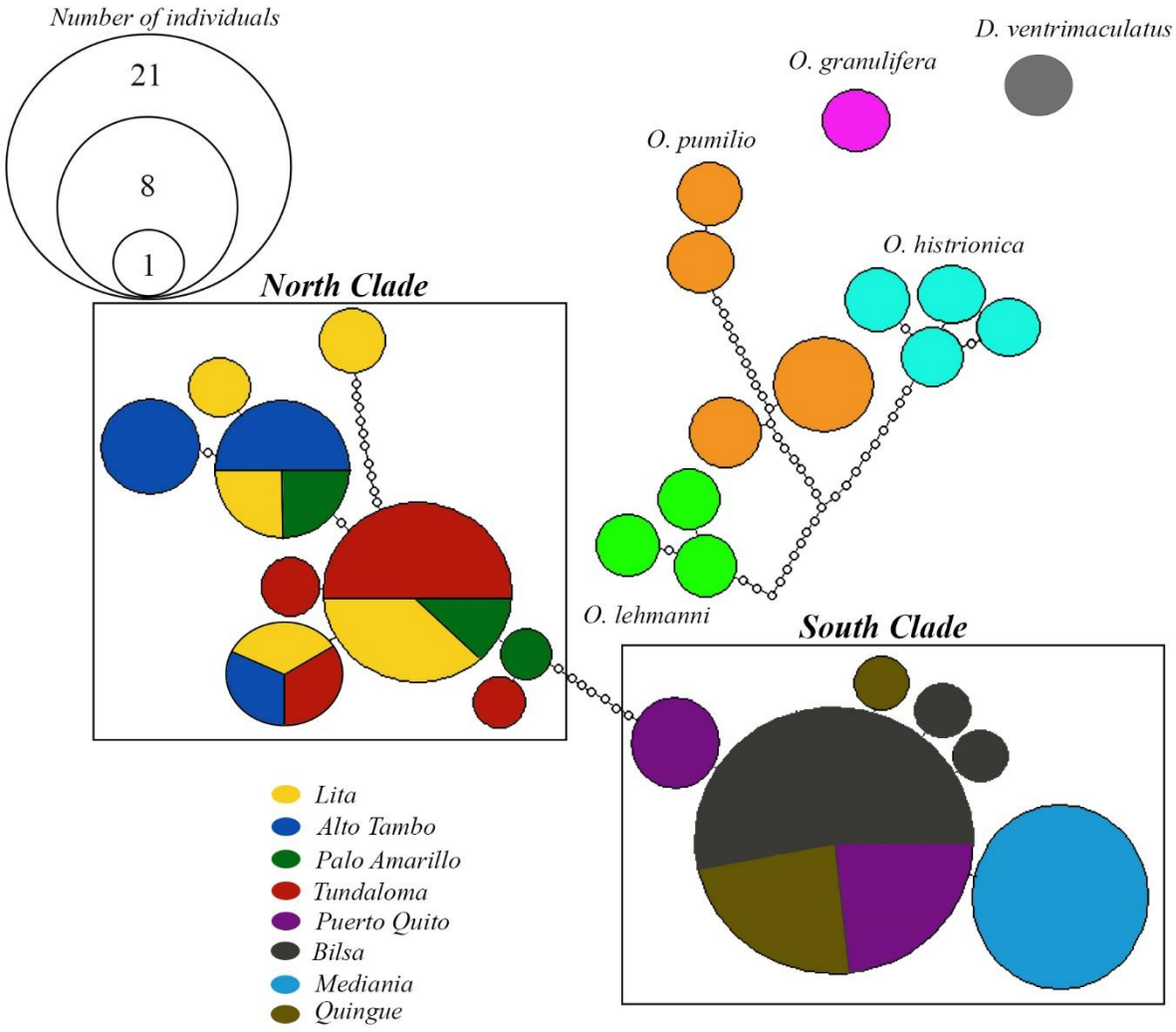

Figure 5. Mitochondrial (COI) haplotype network of individuals of *O. sylvatica*. Sequences of *O. lehmanni*, *O. histrionica*, *O. pumilio* and *O. granulifera* are included as control species. *D. ventrimaculatus* is included as outgroup of the *Oophaga* clade. Circles indicate haplotypes, size is directly proportional to the number of individuals sharing that haplotype. Colours refer to the geographic location of the population that originate haplotype. Pie charts represent the percentage of each population sharing the same haplotype. Little circles in the haplotypes connections indicate the number of mutational steps. Two clades are clearly differentiated within *O. sylvatica* group.

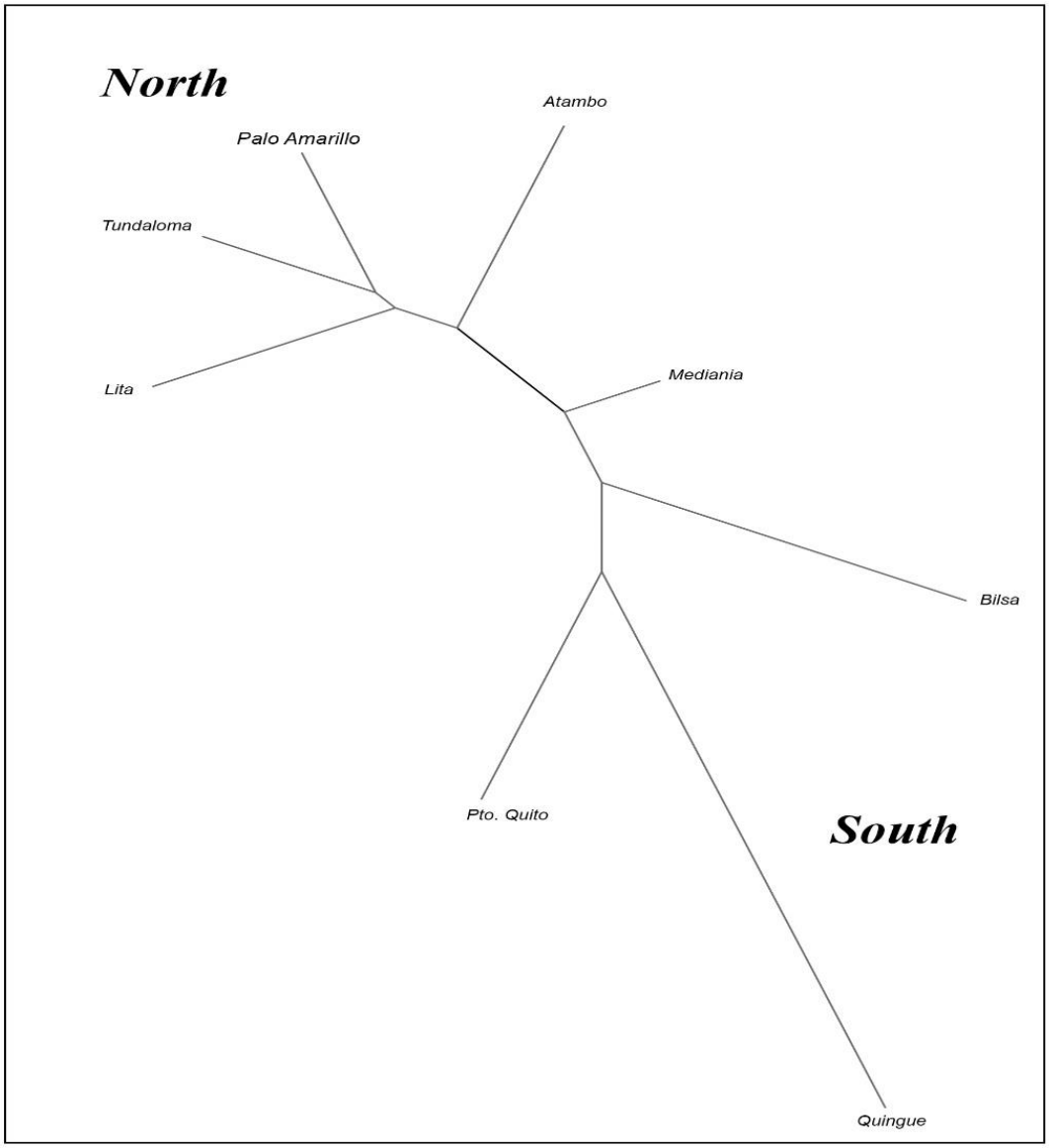

Figure 6. Neighbor-joining tree based in Nei's genetic distances (DS). North and south populations are clearly differenced in two lineages. Mediania population is situated between the two groups.

3.3 Population differentiation

As a starting point to infer the population differentiation I used the clustering algorithm implemented in the program STRUCTURE v2.3.4 (Pritchard et al. 2000; Hubisz et al. 2009). I explored different numbers of populations K to uncover hierarchical population structure. The clear distinction among the two groups when K = 2 is consistent with the hypothesis that all of the populations we studied are divided in two major groups: the northern one including the populations of Lita, Palo Amarillo, Tundaloma and Alto Tambo and the southern one including Bilsa, Quingue, Puerto Quito and Mediania. Further structuring (K= 3) represents divergence of the south western (Bilsa and Quingue) and central-south eastern populations (Puerto Quito and Mediania). Optimal K (Evanno et al. 2005) estimated the most likely number of populations at K = 3 (Figure 7a). The STRUCTURAMA analysis showed K_{NUC}= 2 as the highest probability of the different set of prior values (shape: scale = 2.5:0.52, 2.5:1, 5:2, 10:1, 10:0.1 and 1:1) used for the shape and scale of the gamma distribution of K_{NUC} (Table 4). I further tested the population structure using a multivariate discriminant method approach. Figure 7b illustrates that the entire first DAPC axis clearly separates the northern from the southern populations, while the second one demarcates the Southern populations. Genetically, the central population of Mediania was equidistant from the northern and southern lineages. Based on posterior probabilities of the STUCTURE, all the frogs in this population are of mixed ancestry, with a wide range of individual membership coefficients from the northern and southern lineages.

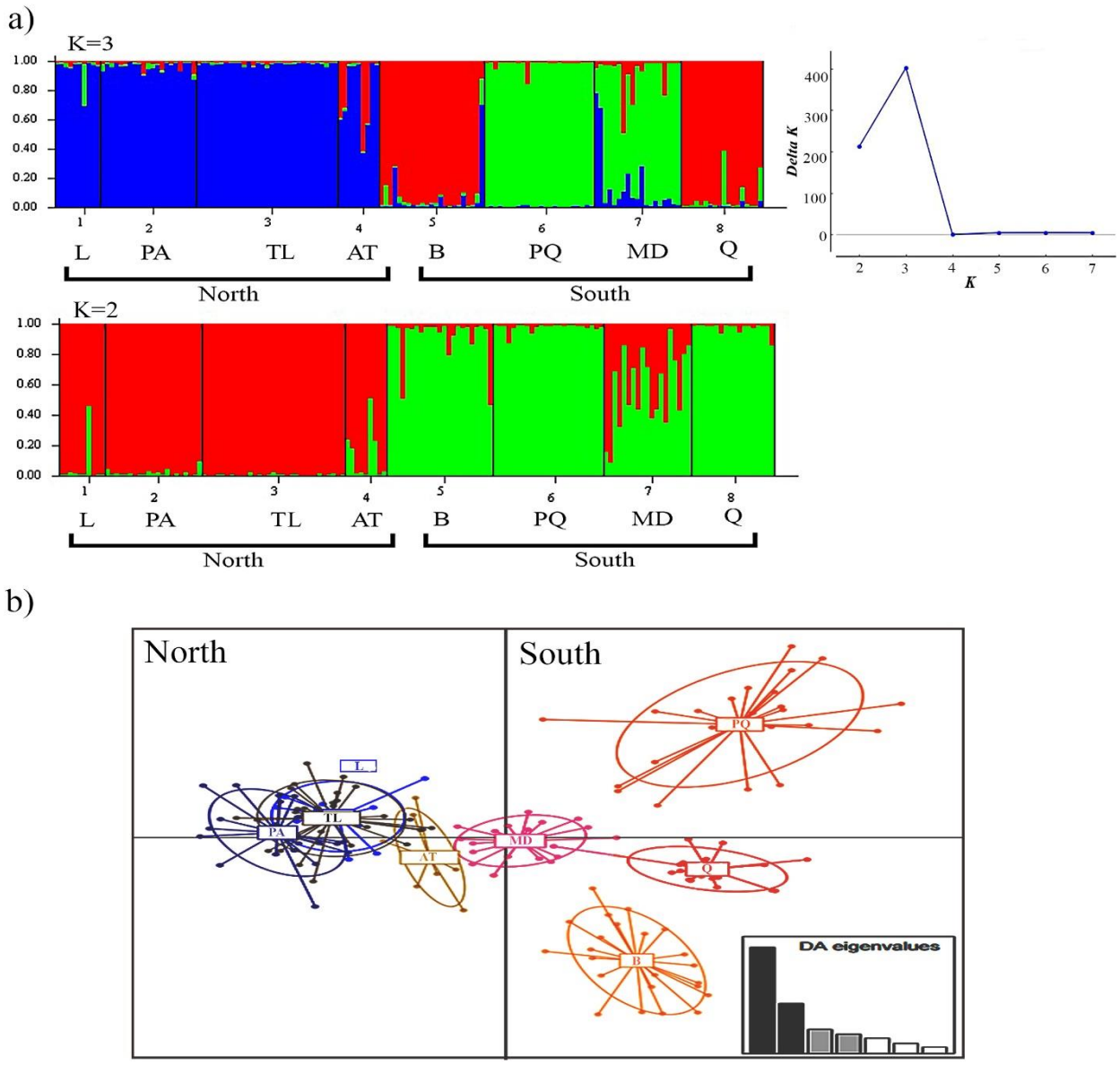

Figure 7. a) Population structure in *O. sylvatica* from microsatellite data. Bayesian clustering shows two (K=2) an three population clusters (K=3), being K=3 the optimal value of K according ΔK method. For K=3, blue color represents the north populations. Red and green the south populations. Some admixture is observable in Mediania population. For K=2 red color represents north populations and green the south ones. Graphic on the right represents the optimal value of K estimated via ΔK method. b) DAPC scatterplot for microsatellite data. The first two components are showed. The first component separates mostly the north and south populations and the second component shows a well differentiated separation between Puerto Quito and Bilsa populations. DAPC eigenvalues axes contribution are showed in the bars graphic. Population codes as in Table 1.

Table 4. Results from STRUCTURAMA analysis at a range of priors for shape and scale in *O. sylvatica* based in microsatellite dataset. The optimal number of inferred populations is shown in light grey

	Priors used in STRUCTURAMA analysis					
Number of populations inferred	Pr [K=i|X] Shape = 2.5 Scale = 0.52	Pr [K=i|X] Shape = 2.5 Scale = 1	Pr [K=i|X] Shape = 5 Scale = 2	Pr [K=i|X] Shape = 10 Scale = 1	Pr [K=i|X] Shape = 10 Scale = 0.1	Pr [K=i|X] Shape = 1 Scale = 1
1	0	0	0	0	0	0
2	0.82	0	0.77	0	0.48	0.91
3	0.17	0.95	0.21	0.85	0.4	0.09
4	0.01	0.05	0.02	0.14	0.11	0
5	0	0	0	0.01	0.01	0
6	0	0	0	0	0	0

3.4 Genetic diversity: microsatellite data

The number of alleles observed across all populations ranged from 36 (*Ooph_13A*) to 10 (*3024*) with an average of 23.46 alleles per locus (Table 5). Observed heterozygosity averaged 0.74 across all populations and 0.66 across all loci ranged from 0.14 (*3024*) to 0.88 (*A3*). In addition, the number of private alleles and its frequency was higher in the south populations (n=59) than north populations (n=47) (Figure 8; Table 5). Furthermore, the inbreeding coefficient (*FIS*) showed that in all populations the values were close to zero, suggesting low probability of consanguinity (Table 5). All sample sites were in Hardy-Weinberg equilibrium and none of them showed statistically significant heterozygote deficits (Table 5). Tests for linkage disequilibrium resulted in significant departures in 4 of 623 possible tests. Significant LD was between two pair of loci *Ooph_109 A/Ooph_44 A* and *Ooph_109 /109.* While the first pair

(*Ooph_109 A/Ooph_44 A*) showed significant LD in only of one of the sampled populations (Tundaloma), the second one showed (*Ooph_109 – 109*) significant LD in the populations of Lita, Palo Amarillo and Tundaloma. Because linkage disequilibrium was only found in certain populations, and because had been based on random linkage disequilibrium (LD) arises by chance each generation in finite populations (particularly if the effective population size –Ne– is small; Laurie-Ahlberg & Weir, 1979; Hill, 1981), I did not exclude these loci from any of subsequent analyses.

The AMOVA for the combined microsatellite loci revealed an overall weak but significant regional (north vs. south) pattern of genetic structure (7%), with significant variation between populations (18%) and very high variation within populations (75%) (Table 6).

Table 5. Summary of nuclear and mitochondrial genetic diversity.

Nuclear DNA						
Populations	Hardy-Weinberg equilibrium (p-value/S.E.)	*mean He*	*FIS*	Number of private alleles	Number of alleles across all populations	
Lita	1.0/0.0000	0.76	0.10	7	*Ooph_110A*	19
					Ooph_109A	31
Tundaloma	1.0/0.0000	0.79	0.14	17	*Ooph_44A*	21
					Ooph_13A	36
P. Amarillo	1.0/0.0000	0.77	0.15	14	*Ooph_6868A*	24
					Ooph_63A	12
Alto Tambo	1.0/0.0000	0.76	0.11	9	*A1*	25
					A2	23
Pto. Quito	1.0/0.0000	0.73	0.09	22	*A3*	32
					A4	35
Bilsa	1.0/0.0000	0.75	0.26	22	*A5*	13
Mediania	0.99/0.0027	0.70	0.07	10	*109*	24
Quingue	0.99/0.0010	0.67	0.07	5	*3024*	10

Mitochondrial DNA					
Lineages	# Haplotypes	Haplotype diversity (Pi)	Nucleotide diversity (Pi)	# Polymorphic sites	# Nucleotide/Parsimony-informative
North	9	0.808	0.00356	22	16
South	6	0.563	0.00118		

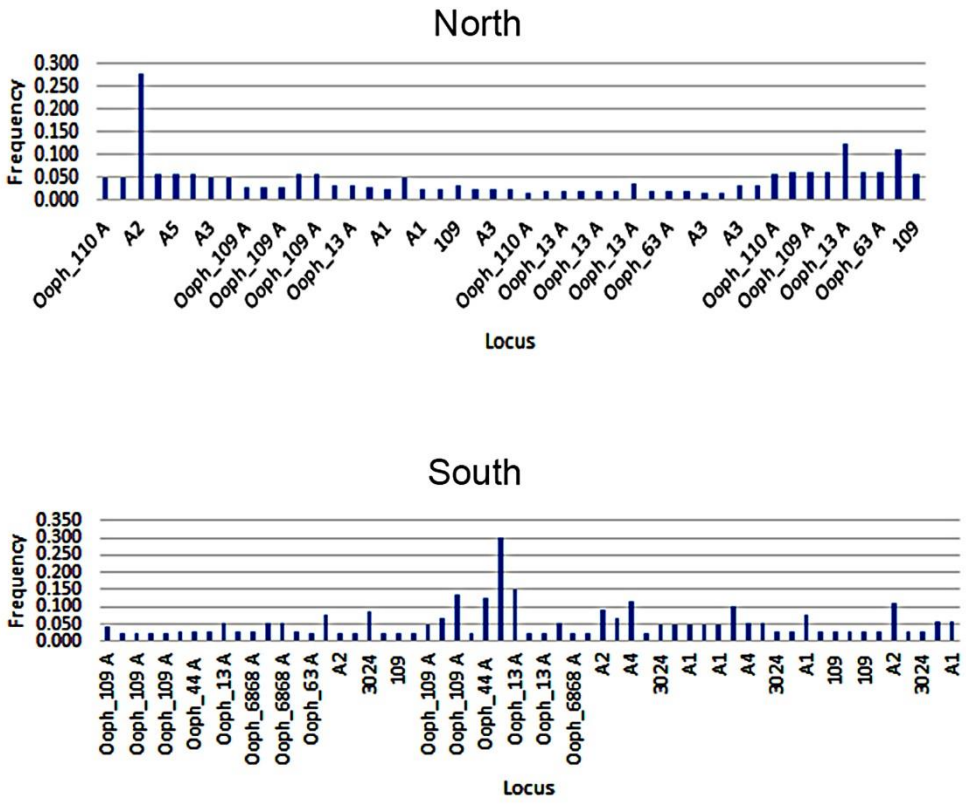

Figure 8. Microsatellite private allelic frequency for both north and south mtDNA lineages. Frequency of private alleles was higher in the south clade.

Table 6. AMOVA analysis results. The molecular variation is mostly retained within populations.

Source	**SS**	**MS**	**Estimate Variation**	**%**
Among Regions	156.867	156.867	1.120	7%
Among Pops	370.909	61.818	2.707	18%
Within Pops	1699.340	11.560	11.560	75%
Total	2227.116		15.387	100%

3.5 Genetic diversity: mitochondrial data

A total of fifteen unique haplotypes were found. Nine in the northern clade and six to the southern one. The haplotype diversity index in the former clade (*Hd*= 0.808) was higher than in the latter one. In the northern clade nine polymorphic sites were found while in the southern one only five polymorphic site were found. Accordingly, the nucleotide diversity in the northern clade (π = 0.00356) was greater than the southern clade (π = 0.00118; Table 5).

Additionally, allelic richness was estimated by rarefaction method to avoid the sample size issue. It measures the allele richness in divergence *(Crd)*, diversity *(Crs)* and the total allelic richness contribution for each population *(Crt)*. In the north, all populations showed a positive allelic richness contribution to their own diversity (*Crs*), but negative contribution to the divergence (*Crd*). Palo Amarillo was the population that most contributed to its allelic richness (*Crs*), In the south, Mediania contributed the most to the total allelic richness (*Crt*), and this was due to its high degree of divergence (*Crd*) from other the populations because in terms of richness within population (*Crs*) was below average. The other populations showed low or negative intrapopulation diversity (*Crs*) and differentiation (*Crd*) contribution (Figure 9).

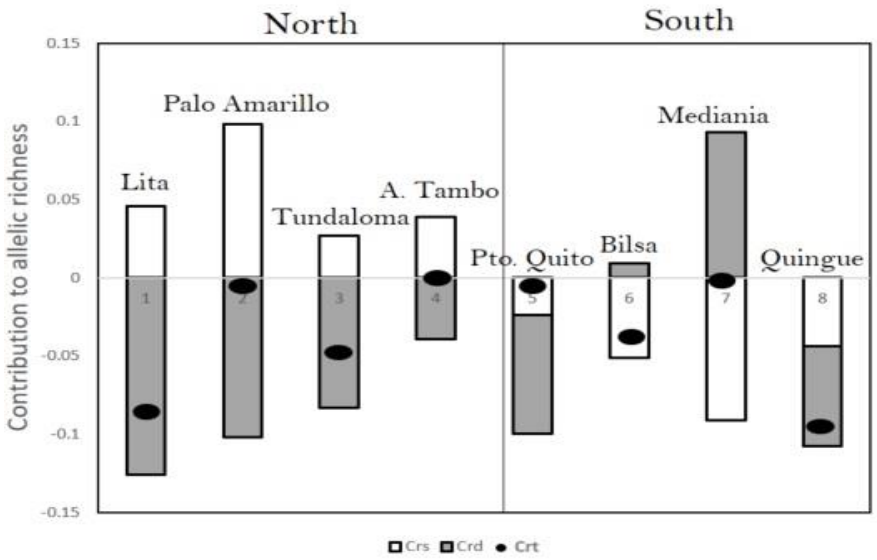

Figure 9. Contribution of mtDNA allelic richness for each population in *O. sylvatica* to the total gene diversity. White bars indicate the intrapopulation diversity contribution (Crs). Grey bars show the contribution to the population differentiation (Crd). Black dots represent the total allelic richness contribution to the diversity (Crt). Palo Amarillo population has the higher contribution to the diversity. Mediania population mostly contributes to the divergence. There is not positive contribution to the total allelic richness.

3.6 Demographic history

For each of the two main mitochondrial lineages, COI sequence data and microsatellite multilocus genotypes were used to determine if the Ecuadorian populations of *O. sylvatica* have undergone past demographic changes. First, the neutrality test showed that the D values for Tajima's D (-0.05212) and Fu's FS (-2.12533) were negative and non-significant ($p>0.05$), causing the rejection of a hypothesis of a sudden population expansion in mtDNA north and south clades. Then, COI was used to determine the mismatch distribution of each mtDNA lineage. The observed distribution of pairwise differences of COI haplotypes between individuals showed an unimodal distribution for both, north and south clades, suggesting a spatial expansion. (Figure 10a). Patterns of historical demography can also be inferred from estimates of the effective population size over time using the EBSP method. It was performed assuming 0.00345 substitutions/site/MY as the expected molecular

evolution rate. North and South populations were considered as independent lineages. The analysis showed that north population have remained constant during the time and just at the end was identified a lightly demographic expansion (between 0.3-0.1 MY). In the south population, the expansion was a little bit higher and was detected between 0.15-0.05 MY (Figure 10b).

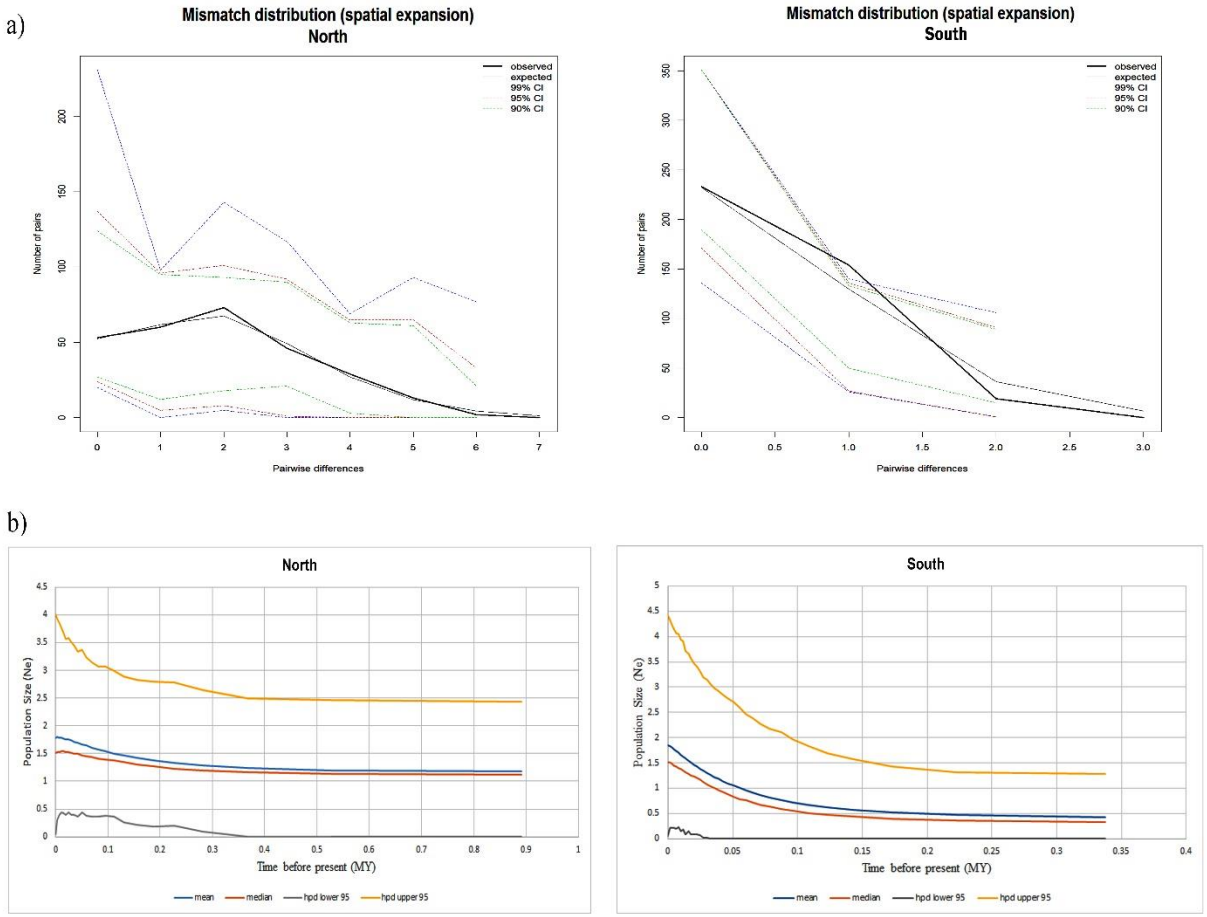

Figure 10. a) Observed distribution of pairwise differences of COI haplotypes between individuals. North and south lineages were analyzed independently. The observed probability show an unimodal shape in both cases (black line). The model is consistent with a recent spatial expansion. b) Extended Bayesian skyline plot for north and south lineages. Both cases show the effective population size has remained constant across time with a light increment in the last 0.2 MY (approx.). Blue and orange line represents the mean and median value of effective population size. Yellow and black lines indicate the upper and lower 95% posterior probability interval.

3.6.1 Bottleneck

Microsatellite dataset was used to determine if populations have suffered bottleneck event which tests recent declines in the effective population size. Table 7 shows the mean observed number of alleles (*mean k*) which was used to calculate the distribution of the expected heterozygosity (*He*) under the assumption of mutation-drift equilibrium (*Heq*) given the sample size and the p values for heterozygosity excess (under TPM mutational model) and deficiency. Overall heterozygosity tests did not reveal signatures of heterozygosity excess and, in consequence, I did not infer recent demographic bottlenecks in any of the populations other than Tundaloma which showed statistically significant number of loci with heterozygosity excess expected under a recent bottleneck ($p < 0.05$). Moreover, for heterozygosity deficiency (*He*<*Heq*), non-significant differences ($p<0.05$) between expected heterozygosity and heterozygosity expected at mutation- drift equilibrium were observed in all populations, suggesting a constant size.

Table 7. Summary of bottleneck analysis based on microsatellite data. Significant values ($p<0.05$) of excess or deficiency of heterozygosity per population are shown in bold.

Population	mean k	mean He	p value 1_tail TPM	p value He<Heq
Lita	8.46	0.80162	0.259277	0.76
P. Amarillo	11.46	0.79741	0.454834	0.57
Tundaloma	11.92	0.80664	**0.000122**	1
Alto Tambo	8.38	0.8047	0.830322	0.19
Bilsa	10.38	0.77476	0.682251	0.34
Pto. Quito	9.46	0.75299	0.526978	0.50
Mediania	5	0.38878	0.65625	0.40
Quingue	4	0.37352	0.65625	0.40

3.7 Environmental clustering and correlation among molecular, phenotypic and environmental variables

The values for bioclimatic variables and the altitude were used to calculate a matrix with Euclidean distances which was used to perform the multivariate and correlation analysis. Hierarchical analysis of bioclimatic variables plus altitude, showed that the studied populations of *O. sylvatica* can be clustered in tree groups: I - Quingue, Tundaloma and Mediania, II- Lita, Palo Amarillo and Alto Tambo and III- Puerto Quito and Bilsa (Figure 11a). A result clustering pattern similar to that observed in the PCA analysis (Figure 11b). The first axis (PCA1) if the latter analysis explains ~84% of the variance and separated the populations by precipitation, with positive values along this axis reflecting higher precipitation regimes. The second axis (PCA2) absorbs ~12% of the variance and discriminated populations based on altitude with high - altitude populations showing positive values (Supplementary 1).

Correlation analyses determined by Mantel test between microsatellite data and environment was significant (p=0.024) suggesting that almost 40% of the molecular variation is explained by the environment. However, when the relationship between molecular and environmental variables controlled by the geographic distances was tested the correlation was not significant (p=0.95). Therefore, the genetic variation is not explained by the environment, it is explicated by the geography. In fact, when the relationship between molecular and geography variables was tested the p value obtained was significant (p=0.001) and showed that the geographic distances explained almost the 75% of the genetic variation (Table 9).

Table 8. The nineteen bioclimatic variables divided in temperature and precipitation

Bioclimatic variables	
Temperature	**Precipitation**
1) Annual Mean Temperature	**12)** Annual Precipitation
2) Mean Monthly Temperature Range	**13)** Precipitation of Wettest Month
3) Isothermality (2/7) (* 100)	**14)** Precipitation of Driest Month
4) Temperature Seasonality (STD * 100)	**15)** Precipitation Seasonality (CV)
5) Max Temperature of Warmest Month	**16)** Precipitation of Wettest Quarter
6) Min Temperature of Coldest Month	**17)** Precipitation of Driest Quarter
7) Temperature Annual Range	**18)** Precipitation of Warmest Quarter
8) Mean Temperature of Wettest Quarter	**19)** Precipitation of Coldest Quarter
9) Mean Temperature of Driest Quarter	
10) Mean Temperature of Warmest Quarter	
11) Mean Temperature of Coldest Quarter	

Table 9. Mantel and partial mantel test between genetic, environmental and geographic variables.

Variables	**Pearson correlation (r)**	**Significance (p<0.05)**
Genetic vs Environment	0.380	0.024
Genetic vs Geography	0.746	0.001
Genetic vs Environment controlled by Geography	-0.204	0.905

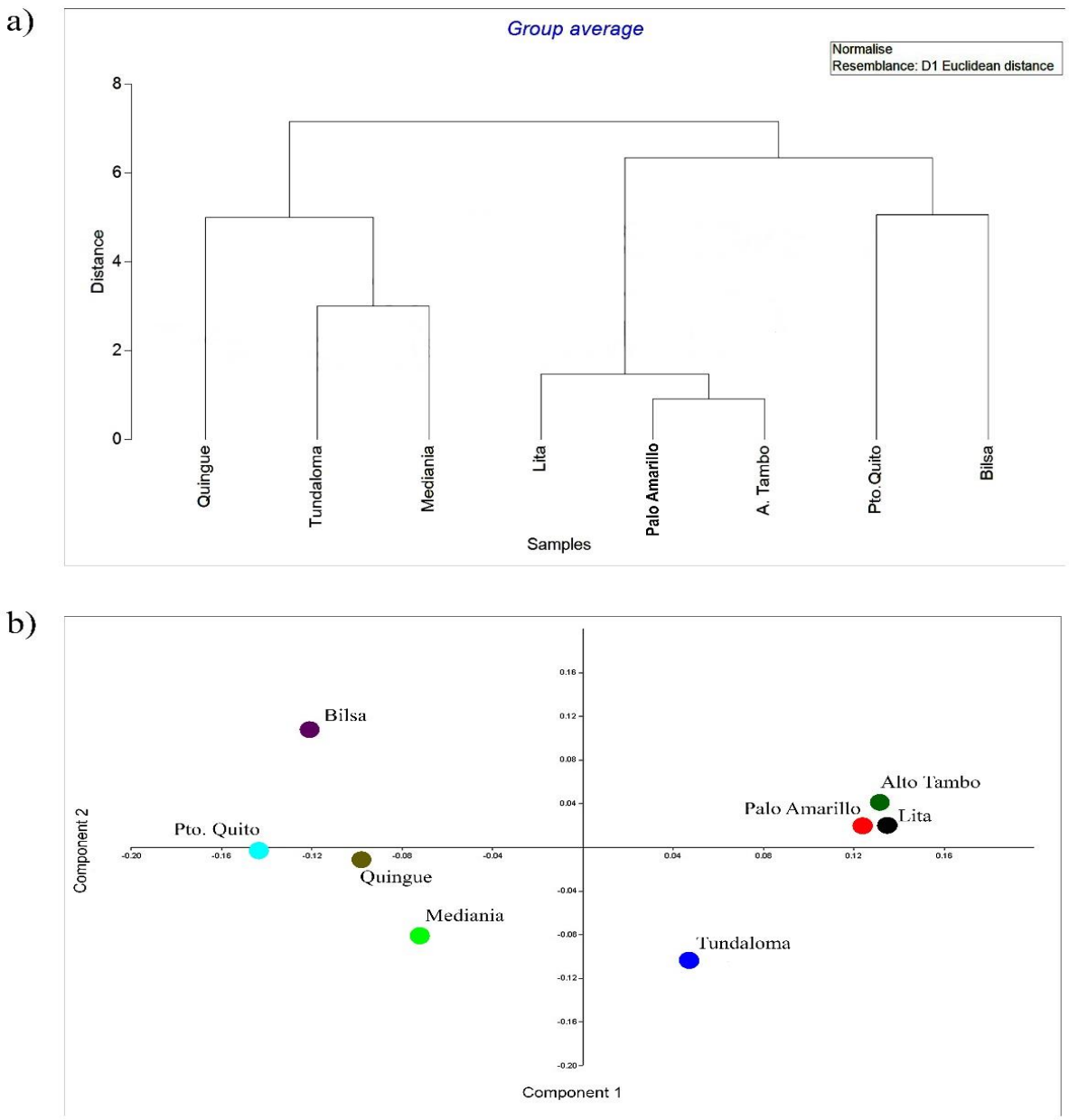

Figure 11. a) Cluster of environmental variables based on Euclidean distances. Localities are grouped in three environmental clusters. First cluster: Quingue, Tundaloma and Mediania. Second cluster: Lita, Palo Amarillo and Alto Tambo. Third cluster: Pto. Quito and Bilsa. b) PCA of environmental variables. Dots represents each locality. Precipitation and altitude mostly retained the variation (PC1= 84%, PC2=12%).

CHAPTER 4

DISCUSSION

4.1 Phenotypic differentiation

One of my main objectives was to quantify the phenotypic diversity in *O. sylvatica* and examine the extent to which its spatial distribution might be concordant with genetic and ecological data. My results provide evidence that in *O. sylvatica* geographically dispersed phenotypes are shared among phylogeographically structured populations. Different evolutionary processes can account for such a mismatch between genotypic and phenotypic spatial variation (reviewed in Zamudio et al. 2016). Parallel adaptation (*i.e.* the independent evolution of similar adaptive phenotypes in genetically isolated lineages) is a likely mechanism for shared phenotypic variation in species with strong phylogeographic structure. A recent of the diversification and convergence of coloration among lineages of aposematic Harlequin poison frogs (*O. histrionica* complex) conclude dark aposematic phenotypes, showing strong color and/or luminance contrast, have evolved independently at least twice in this species complex, and that diversification of dorsal background coloration in this group predates that of the different mitochondrial lineages (Posso & Andres, 2017). Thus, it is possible that similar coloration phenotypes have evolved independently in the different *O. sylvatica* lineages. However, this is not the only possible explanation. Rapid phenotypic diversification often involves complex histories of hybridization (gene flow) that leave conflicting phylogenetic signatures across the genome. Although most of the alleles crossing lineage boundaries are likely to be neutral or deleterious in their new genetic background, the adaptive introgression of genetic material is possible if the gained alleles confer a selective advantage to the recipient species (Allendorf et al. 2012). For instance, a *Vkorc1* allele

associated with resistance to the rodenticide warfarin has “jumped” across different mouse species (Song et al. 2011). Similarly, in *Heliconious*, a group of neotropical butterflies that display Müllerian mimicry, alleles for red wing coloration in clade share similar wing coloration patterns due to introgression between species (Zhang et al. 2016).

Based on COI data, it appears that *O. sylvatica* consists of at least a Northern and a Southern lineage that overlap the Mache-Chindul natural protected area which is part of the Esmeraldas river basin. It has been recently identified as an endemic area in western South America and represents an important portion of the last remnants of coastal tropical forest on northwestern Ecuador (Ortega-Andrade et al. 2010). Members of the populations sampled in this area showed admixed nuclear genomes. Thus, it is possible that observed spatial mismatch is the result of adaptive introgression of coloration alleles from one mitochondrial clade to another. An alternative explanation, is the genotypic/phenotypic mismatch is due to the retention of ancestral polymorphism which may be either neutral, or more likely maintained by stabilizing (frequency or density dependent) selection. Clearly, further genetic analyses are necessary to distinguishing between these hypotheses.

4.2 Genetic differentiation

Amphibians in general are poor dispersers and highly phylopatric, with strict habitat specificity and physiological requirements (reviewed by Hauswaldt et al. 2011). Consequently, significant population genetic structure is often expected, especially over moderately large geographic distances. At a large geographical scale, Bayesian clustering, F_{ST} pairwise comparisons and AMOVA analyses all rejected the null hypothesis of panmixia, supporting the existence of different groups of genetic diversity among Ecuadorian regions. The Bayesian clustering analysis suggests the existence of at least two main genetic units of

Diablito poison frogs concordant with that of the mitochondrial clades (N + S), and most likely three (N, SE, and SW) genetic clusters. At a more local scale, Northern populations were relatively more homogeneous than the Southern ones. A result that might be explain by the geographical proximity of the sampled populations in the North. The DAPC plotting analysis also supported this pattern.

Microsatellite markers revealed that the majority of the variation was found within populations rather than among populations (or between clades). High intrapopulation variability is often explained as either the result high levels of gene flow or the fact that populations have not been separated long enough to accumulate detectable genetic differences (Allendorf et al. 2012). However, interpreting microsatellite data correctly is not necessarily straight forward. Due to their high levels of polymorphism and the relatively small sample sizes often employed in population genetic studies (e.g., 30 individuals per population) the estimation of allele frequencies may not be accurate (Allendorf et al. 2012). Also, homoplasy caused by mutation may influence population structure inferences because it depresses gene diversity which, in turn, may lead to underestimation of population differentiation. Similarly, measures of population subdivision such as F_{ST} (Wright, 1951) and R_{ST} (Slatkin, 1995) may be poor measures of genetic differentiation when the level of diversity (and heterozygosity) is high (Charlesworth, 1998). Thus, the relatively high level of intrapopulation variability found in my study may not be a biological phenomena associated but a consequence of the number and class of genetic markers.

4.3 Effective population size and population size change

Past population size changes were inferred for both mitochondrial populations and in the microsatellite data. Similar patterns were inferred from the two marker types. Analyses of microsatellite data suggest that populations are relatively stable, while EBSP analyses of mitochondrial data suggest this stability was preceded by population growth back in time. The time of stronger population expansion reconstructed from mitochondrial data roughly corresponds to that of the late Pleistocene at the beginning of the present interglacial period (~ 20.000 years ago). There is ample evidence that the biota of the temperate zones has been shaped by the Pleistocene glacial-interglacial periods (Brown & Knowles, 2012). Similarly, the flora and fauna of tropics have been also affected by Pleistocene climatic shifts. In South America, a series of humid-arid cycles drastically and repeatedly altered vegetation patterns during the Quaternary (Enrst & O'Hara, 1986; Costa, 2003). Periods of reduced rainfall, led to the temporary fractionation of previously continuous rainforests into forest islands separated by wide belts of savanna. The biota of the tropical lowland forests was isolated into these forest refuges and given the opportunity to diverge. The approximate time of divergence between the north and south clade in *O. sylvatica* was around 1.2 MY ago, this matches with the Günz glaciation (~ 1.1 MY). I hypothesized that North and South populations became isolated during this glaciation period and with the return of more mesic conditions, a secondary contact (hybridization) between the two mitochondrial lineages found in *O. sylvatica* occurred during the period of demographic expansion which showed a stronger signal ~ 15.000 years ago, at the beginning of the present interglacial period (Supplementary 2) . This situation could explain the admixture observed in the population of Mediania, geographically localized in the middle between North and South populations.

4.4 Conservation and management

Evolutionarily significant units (population units that merits separate management and has a high priority for conservation) are defined by two components that are not necessarily correlated: adaptive distinctiveness and historical isolation. Adaptive differences can certainly arise and become fixed in the presence of gene flow (Allendorf et al. 2012) and reproductive isolation does not *per se* guarantee the existence of adaptive differences among isolated lineages, particularly in homogeneous environments. However, because adaptive distinctiveness in not always easy to estimate, efforts to define ESUs have often emphasized reproductive isolation rather than the maintenance of adaptive differences (Crandall et al. 2000). As a result, the genetic value of populations has been based, essentially, on a measure of the evolutionary isolation of each relative to all others (genetic uniqueness). My genetic analyses resulted in estimates of how divergent or "atypical" each population were from the "average" diablito frog population in terms of microsatellite and mtDNA, and hence, population values in terms of unique genetic variation. According this criterion, each of the *O. sylvatica* monophyletic clade should be treated separately. In the Northern clade, the population of Palo Amarillo should be probably prioritized since it is the population that contributes the most to haplotype richness (mtDNA) and there are no significant differences in genetic uniqueness among populations (average pairwise Nei's distance range: 0.244-0.284). In the Southern clade, the population of Mediania showed a strong signal of genetic uniqueness (Crd) associated with a private mitochondrial haplotype that may be fixed in this population. However, mitochondrial uniqueness was not associated with that at the nuclear genome. According to the microsatellite dataset, the population of Bilsa showed the strongest signal isolation. These results imply in the evaluation of population distinctiveness a single type of genetic marker is problematic. Given these contrasting data, one could further

prioritize amongst populations based on other evolutionary and/or ecological characteristics. The population of Mediana is geographically located in the center of the *O. sylvatica* distribution in Ecuador, and shows signs of genetic admixture between the Northern and Southern clades. If the ultimate conservation goal be to conserve ecological and evolutionary processes (rather than to the products of those processes; Morizt, 1994), we need to be looking at all populations and see what is meaningful across the whole species rather than just looking at individual populations. One way to implement this holistic approach is to look at the patterns of genetic and ecological exchangeability (*sensu* Crandrall et al. 2000). Individuals from different populations are genetically exchangeable if there is ample gene flow between populations. Similarly, individual from different populations are ecologically exchangeable if they can be moved between populations, and can occupy the same ecological niche or selective regime. Both ecological and genetic exchangeabilities can be partitioned along a time scale to generate a continuum of management options as oppose to just two categories (ESU or not, Cradall et al 2000). In *O. sylvatica*, historical genetic exchangeability can only be supported with confidence between the Northern and Southern clades. The low ($F_{st} < 0.05$) to moderate ($F_{st} = 0.05$-0.15, Hartl & Clark, 1997) population structure detected by microsatellites values of population differentiation suggesting recent genetic exchange.

As other members of the *Oophaga* genus the Diablito poison frogs show a remarkable variation in coloration. The general consensus is that such phenotypic diversification is not the result of genetic drift but a combination of natural and sexual selection (e.g. Summers et al. 1997; Hagemann & Pröhl, 2007; Rudh et al. 2007; Brown et al. 2010; Wang & Summers, 2010; Medina et al. 2013; Richards-Zawacki, et al. 2012). Therefore, coloration in *Oophaga* frogs may be used as a proxy for the ecological exchangeability. My multivariate analyses clearly reject the null hypothesis of ecological uniformity across populations. This combination of molecular and morphological exchangeabilities suggests that *O. sylvatica*

should be treated as as two distinct populations with recent admixture and loss of genetic distinctiveness (Case 3 in Crandall et al. 2000).

4.5 Conclusions and future directions

My study revealed that *O. sylvatica* encompasses at least two mitochondrial clades with significant intraspecific divergence in aposematic coloration patterns. Based on both genetic and ecological exchangeability criteria, I propose that the Ecuadorian population of this species should be managed as two different units. Nevertheless, this recommendation should be taken with care given the limited number of individuals and populations sampled in this study. Genetic data suggest the possible existence of a highly divergent mitochondrial lineage in the north. Therefore, a better survey is necessary paying special to the populations located in southern Colombia and northern Ecuador, along the upper Rio Mira basin, which may represent a dispersal barrier. It is also critical to sample the central range of the distribution where allelic admixture and unique haplotypes were observed. New analyses using a higher number of molecular markers (possibly SNPs) as well as a broader range of phenotypic and ecological traits (mating calls, bioclimatic niches, predator communities, etc.) are also needed to better estimate the appropriate (IUCN) conservation status and conservation units in this charismatic species.

LIST OF REFERENCES

Allendorf, F. W., Luikart, G. & Aitken, S. (2012). Conservation and the Genetics of Populations. Second Edition. Wiley-Blackwell, Cambridge, Massachusetts, USA.

AmphibiaWeb. (2016). <http://amphibiaweb.org> University of California, Berkeley, CA, USA. Accessed 12 Dec 2016.

Arteaga, A., Pyron, R.A., Peñafiel, N., Romero-Barreto, P., Culebras, J. & Bustamante, L. (2016). Comparative Phylogeography Reveals Cryptic Diversity and Repeated Patterns of Cladogenesis for Amphibians and Reptiles in Northwestern Ecuador. *PLoS ONE,* **11**, 4.

Baillie, J.E.M., Hilton-Taylor, C. & Stuart, S. (2004). 2004 IUCN Red List of Threatened Species: A global species assessment. In, p. 217. IUCN, Gland, Switzerland and Cambridge, UK

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A. & Drummond, A.J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Comput*, **Biol.**, 10.

Brown, J.L. & Knowles, L.L. (2012). Spatially explicit models of dynamic histories: examination of the genetic consequences of Pleistocene glaciation and recent climate change on the American Pika. *Mol Ecol,* **21(15),** 3757–3775.

Brown, J.L., Maan, M.E., Cummings, M.E. & Summers, K. (2010). Evidence for selection on coloration in a Panamanian poison frog: a coalescent-based approach. *J Biogeogr,* **37(5),** 891–901.

Charlesworth, B. (1998). Measures of divergence between populations and the effect of forces that reduce variability. *Molecular Biology and Evolution,* **15,** 538–543.

Clarke, K.R. & Gorley, R.N. (2015). PRIMER v7: User Manual/Tutorial. PRIMER-E. *Plymouth,* 296pp.

Clement, M., Posada, D. & Crandall, K.A. (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9,** 1657–1659

Costa, L.P. (2003). The historical bridge between the Amazon and the Atlantic Forest of Brazil: a study of molecular phylogeography with small mammals. *Journal of Biogeography,* **30,** 71– 86.

Coloma, L. A., Ron, S., Grant, T. & Lötters, S. (2004). *Oophaga sylvatica*. The IUCN Red List of Threatened Species 2004: e.T55203A11264944. http://dx.doi.org/10.2305/IUCN.UK.2004.RLTS.T55203A11264944.en. Downloaded on 15 December 2016.

Corander, J., Sirén, J. & Arjas, E. (2008). Bayesian spatial modelling of genetic population structure. *Comp Stat,* **23,** 111–129.

Cornuet, J.M. & Luikart, G. (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics,* **144,** 2001–2014.

Crandall, K. A., Bininda-Emonds, O. R. P., Mace, G. M. & Wayne, R. (2000). Considering evolutionary processes in conservation biology. *Trends in Ecology & Evolution* **15,** 290–295.

Crawford, A. (2003) Relative rates of nucleotide substitution in frogs. *Journal of Molecular Evolution*, **57**, 636-641.

Daszak, P., Berger, L., Cunningham, A., Longcore, J., Brown, C. & Porter, D. (2004). Experimental evidence that the bullfrog (Rana catesbeiana) is a potential carrier of chytridiomycosis, an emerging fungal disease of amphibians. *Herpetological Journal*, **14,** 201-207.

Dizon, A. E., Lockyer, W. F., Pemn, D. P., Demaster, J. & Sisson, J. (1992). Rethinking the stock concept: a phylogeographic approach. *Conservation Biology,* **6,** 24-36.
Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. (2012). Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol*, **29,** 1969-1973.
Duellman, W. E. & Trueb, L. (1986). Biology of Amphibians. McGraw-Hill, New York.
Earl, D.A. (2011). Structure Harvester v0.6.5 Available at http://users.soe.ucsc.edu/*dearl/software/structureHarvester/
Ernst, M. & O' Hara, R. (1986). The Biogeografic Evidence Supporting the Pleistocene Forest Refuge Hypothesis. *Evolution,* **40(1),** 55-67.
Excoffier, L., Laval, G. & Schneider, S. (2005). Arlequin version 3.5: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online,* **1,** 47–60.
Frankham, R., Briscoe, D. A. & Ballou, J. D. (2002). Introduction to conservation genetics. Cambridge University Press, New York, New York, USA.
Funkhouser, J. (1956) New frogs from Ecuador and southwestern Colombia. *Zoologica*, **41**, 73-80.
Gehara, M., Summers, K. & Brown J.L. (2013). Population expansion, isolation and selection: novel insights on the evolution of color diversity in the strawberry poison frog. *Evolutionary Ecology,* **27,** 797–824.
Grant, T., Frost, D.R., Caldwell, J.P., Gagliardo, R.O.N., Haddad, C.F.B., Kok, P.J.R., Means, D.B., Noonan, B.P., Schargel, W.E. & Wheeler, W.C. (2006) Phylogenetic systematics of dart-poison frogs and their relatives (Amphibia: Athesphatanura: Dendrobatidae). *Bulletin of the American Museum of Natural History*, 1-262.
Guillot, G., Mortier, F. & Estoup, A. (2005). GENELAND: a computer package for landscape genetics. *Molecular Ecology,* **5(3),** 712-715.
Hagemann, S. & Pröhl, H. (2007). Mitochondrial paraphyly in a polymorphic poison frog species (Dendrobatidae; *D. pumilio*). *Mol Phylogenet Evol*, **45,** 740–747.
Hammer, Ø., Harper, D. & Ryan, P. (2001). PAST: Paleontological Statistics Software Package for education and data analysis. *Palaeontolia Electronica,* **4(1),** 9.
Hartl, D.L. & Clark, A.G. (1997). Principles of Population Genetics, 3nd edn. Sinauer Associates, Inc, Sunderland, MA.
Hauswaldt, J., Ludewig, A.K., Vences, M. & Pröhl, H. (2011). Widespread co-occurrence of divergent mitochondrial haplotype lineages in a Central American species of poison frog (*Oophaga pumilio*). *J. Biogeogr*, **38,** 711-726.
Hill, W.G. (1981). Estimation of effective population size from data on linkage disequilibrium. *Genet. Res,* **38,** 209– 216.
Ho, S.Y.W. & Shapiro, B. (2011). Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology Resources,* **11,** 423–434.
Hostetler, J., Onorato, D., Jansen, D. & Oli, M. (2012). A cat's tale: the impact of genetic restoration on Florida panther population dynamics and persistence. *Journal of Animal Ecology,* **82(3),** 1365-2656.
Hubisz, M. J., Falush, D., Stephens, M. & Pritchard J.K. (2009). Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Resour,* **9,** 1322-1332.
Huelsenbeck, J.P., Andolfatto, P. & Huelsenbeck, E.T. (2011). Structurama: Bayesian Inference of Population Structure. *Evol Bioinform Online*, **7,** 55-59.
IUCN (2016) *IUCN Red List of Threatened Species. Version 2012.1*,. Available at: http://www.iucnredlist.org. (accessed March 2017).

Johnson, W., Onorato, D., Roelke, M., Land, D., Cunningham, M., Belden, R., McBride, R., Jansen, D., Lotz, M., Shindle, D., Howard, J., Wildt, D., Penfold, L., Hostetler, J., Oli, M. & O'Brien, S. (2010). Genetic Restoration of the Florida Panther. *Science,* **329 (5999),** 1641-1645.

Jombart, T., Devillard, S. & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC genetics*, **11,** 94.

Lacy, R. (1997). Importance of genetic variation to the viability of mamalian populations. *Journal of Mammalogy*, **78,** 320-335.

Laurie-Ahlberg, C.C. & Weir, B.S. (1979). Allozymic Variation and Linkage Disequilibrium in Some Laboratory Populations of DROSOPHILA MELANOGASTER. *Genetics,* **4,** 1295–1314.

Liberg, O., Andrén, H., Pedersen, H.C., Sand, H., Sejberg, D., Wabakken, P., Åkesson, M. & Bensch, S. (2005). Severe inbreeding depression in a wild wolf *Canis lupus* population. *Biology Letters*, **1,** 17–20.

Librado, P. & Rozas, J. (2009). DnaSP v5: A Software for Comprehensive Analysis of DNA Polymorphic Data. *Bioinformatics*, **25,** 1451-1452.

Lips, K. R., Diffendorfer, J., Mendelson, J. R. & Sears, M. W. (2008). Riding the wave: reconciling the roles of disease and climate change in amphibian declines. *PLoS Biology*, **6,** 441-454.

Lötters, S., Glaw, F., Kohler, F. & Castro, F. (1999). On the geographic variation of the advertisement call of *Dendrobates histrionicus* Berthold, 1845 and related forms from north-western South America. *Herpetozoa*, **12**, 23-38.

Maehr, D. S. & Caddick, G. B. (1995). Demographics and Genetics Introgression in the Florida Panther. *Conservation Biology,* **9(5),** 1523-1739.

McGugan, J.R., Byrd, G.D., Roland, A.B., Caty, S.N., Kabir, N., Tapia, E.E., Trauger, S.A., Coloma, L.A. & O' Connell, L.A. (2016). Ant and Mite Diversity Drives Toxin Variation in the Little Devil Poison Frog. *Journal of Chemical Ecology*, **42(6),** 537-551.

Medina, I., Wang, I.J., Salazar, C. & Amézquita, A. (2013). Hybridization promotes color polymorphism in the aposematic harlequin poison frog, *Oophaga histrionica*. *Ecol. Evol,* **3,** 4388-4400.

Meffe, G. K. & Carroll, C. R. (1997). Principles of Conservation Biology, 2nd Edition. Sinauer Associates, Sunderland, MA.

Moritz, C. (1994). Defining 'evolutionary significant units' for conservation. *Trend in Ecology and Evolution*, **9,** 373-375.

Nei, M. (1972). Genetic distance between populations. *American Naturalist,* **106,** 283–292.

Olson, D.H., Aanensen, D.M., Ronnenberg, K.L., Powell, C.I. & Walker, S.F. (2013). Mapping the global emergence of *Batrachochytrium dendrobatidis*, the amphibian chytrid fungus. *PLoS ONE*, **8(2),** e56802. doi:10.1371/journal.pone.0056802.

Oksanen, J., Blanchet, F., Kindt, R., Legendre, P. & Minchin, P. (2013). vegan: Community Ecology Package. R package version 2.0-6. Available at: http://mirror.bjtu.edu.cn/cran/web/packages/vegan/

Ortega-Andrade, M., Bermingham, J., Aulestia, C. & Paucar, C. (2010). Herpetofauna of the Bilsa Biological Station, province of Esmeraldas, Ecuador. *Check List,* **1(6),** 119-154.

Ortego, J.I.N., Calabuig, G., Cordero, P.J. & Aparicio, J.M. (2007). Egg production and individual genetic diversity in lesser kestrels. *Molecular Ecology*, **16,** 2383–2392.

Ortiz, D. A., Coloma, L. A., Frenkel, C. & Pazmiño-Armijos, G. (2017). *Oophaga sylvatica*. En: Ron, S. R., Guayasamin, J. M., Yanez-Muñoz, M. H., Merino-Viteri,

A., Ortiz, D. A. & Nicolalde, D. A. (2016). AmphibiaWebEcuador. Version 2016.0. Museo de Zoología, Pontificia Universidad Católica del Ecuador.<http://zoologia.puce.edu.ec/vertebrados/anfibios/FichaEspecie.aspx?Id=1261>, accessed May 22, 2017.

Peakall, R. & Smouse, P.E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics,* **28,** 2537-2539.

Petit, R. J., Mousadik, A. & Pons, O. (1998). Identifying populations for conservation on the basis of genetic markers. *Conserv. Biol,* **12,** 844–855.

Phillott, A.D., McDonald, K.R. & Skerratt, L.F. (2011) Inflammation in digits of unmarked and toe-tipped wild hylids. *Wildlife Research*, **38**, 204-207.

Piry, S., Luikart, G. & Cornuet, J.M. (1999). BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity,* **90,** 502–503.

Posada, D. (2008). jModelTest: Phylogenetic Model Averaging. *Mol. Biol. Evol*, **25,** 1253-1256.

Posso-Terranova, A. & Andres, J.A. (2017). Diversification and convergence of aposematic phenotypes: truncated receptors and cellular arrangements mediate rapid evolution of coloration in Harlequin poison frogs. *Evolution,* submitted.

Posso-Terranova, A. & Andrés, J.A. (2016a). Ecology, molecules and colour: Multivariate species delimitation and conservation of Harlequin poison frogs.

Posso-Terranova A. & Andrés, J.A. (2016b). Complex niche divergence underlies lineage diversification in *Oophaga* poison frogs. *Journal of Biogeography.*

Primmer, C.R. (2009). From Conservation Genetics to Conservation Genomics. *Annals of the New York Academy of Science,* **1162,** 357-368.

Pritchard, J.K., Stephens, M. & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics,* **155,** 945-959.

R Development Core Team. (2011). R: A Language and Environment for Statistical Computing. Vienna, Austria: *The R Foundation for Statistical Computing*, ISBN: 3-900051-07-0. Available online at http://www.R-project.org/.

Ray, N., Currat, M. & Excoffier, L. (2003). Intra-Deme Molecular Diversity in Spatially Expanding Populations. *Mol. Biol. Evol.,* **20(1),** 76-86.

Raymond, M. & Rousset, F. (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Heredity,* **86,** 248–249.

Reeves, P.A. & Richards, C.M. (2009). Accurate Inference of Subtle Population Structure (and Other Genetic Discontinuities) Using Principal Coordinates. *PLoS ONE,* **4(1)**: 4269.

Richards-Zawacki, C.L., Wang, I.J. & Summers, K.S. (2012). Mate choice and the genetic basis for colour variation in a polymorphic dart frog: inferences from a wild pedigree. Mol Ecol, **21,** 3879–3892.

Rogers, A. R. & Harpending, H. C. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.,* **9,** 552-569.

Roland, A., Santos, J. Carriker, B., Caty, S., Tapia, E., Coloma, L. & O'Connell, L. (2016). Radiation and hybridization of the Little Devil poison frog (*Oophaga sylvatica)* in Ecuador. doi: https://doi.org/10.1101/072181

Rudh, A., Rogell, B. & Hoglund, J. (2007). Non-gradual variation in colour morphs of the Strawberry Poison Frog *Dendrobates pumilio*: genetic and geographical isolation suggest a role for selection in maintaining polymorphism. *Mol Ecol,* **16,** 4284–4294.

Ryder, O. A. (1986). Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology and Evolution,* **1,** 9-10.
Santos J.C., Coloma, L.A. & Summers, K. (2009). Amazonian Amphibian Diversity Is Primarily Derived from Late Miocene Andean Lineages (C Moritz, Ed,). *PLoS Biology,* **7,** e1000056.
Schneider, S. & Excoffier, L. (1999). Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **152**, 1079–1089
Sheffield, J. (2007). ImageJ: A useful tool for biological image processing and analysis. *Microsc. Microanal*. **13,** 200-201.
Silverstone, P.A. (1975). A revision of the poison-arrow frogs of the genus *Dendrobates* Wagler. *Natural History Museum of Los Angeles County*, **21**, 1-55.
Slatkin, M. & Hudson, R. (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics,* **129**, 555-562.
Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics,* **139,** 457–462.
Song, Y., Endepols, S., Klemann, N., Richter, D., Matuschka, F-R. & Shi, C-H. (2011). Adaptive introgression of anticoagulant rodent poison resistance between old world mice. *Curr Biol,* **21(15),** 1296–1301.
Stevens, M., Parraga, C.A., Cuthill, I.C., Partridge, J.C. & Troscianko, T.S. (2007) Using digital photography to study animal coloration. *Biological Journal of the Linnean Society*, **90**, 211-237
Stamatakis, A. (2014). RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics,* **10,** 1093/bioinformatics/btu033
Stuart, S. N. (2004). Status and trends of amphibian declines and extinctions worldwide. *Science,* **306,** 1783 – 1786.
Summers, K., Bermingham, E., Weight, L., Mc Cafferty, S. & Dahlstrom, L. (1997). Phenotypic and genetic divergence in three species of dart poison frogs with contrasting parental behaviour. *J. Hered*, **88,** 1.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*, **22,** 4673-4680.
United States Government. (1988). Endangered Species Act of 1973, as amended through the 100th Congress. U.S. Department of the Interior.
Vuilleumier, B. S. (1971). Pleistocene changes in the fauna and flora of South America. *Science,* **173,** 771-780.
Wake, D. B. & Vredenburg, V.T. (2008). Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proceedings of the National Academy of Sciences*, **105,** 11466–11473.
Wang, I.J. & Summers, K. (2010). Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. *Mol Ecol,* **19(3),** 447–458.
Waples, R.S. (1995). Evolutionary significant units and the conservation of biological diversity under the endangered species act. *American Fisheries Society Symposium*, **17,** 8-27.
Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics,* **15,** 323–354.

Zamudio, K.R., Bell, R.C, & Mason, N.A. (2016). Phenotypes in phylogeography: Species traits, environmental variation, and vertebrate diversification. *PNAS*, **113(29),** 8041-8048. doi: 10.1073/pnas.1602237113

Zhang, J., Chiodini, R., Badr, A.& Zhang, G. (2011). The impact of next-generation sequencing on genomics. *Journal of Genetics and Genomics*, **38,** 95–109.

Zhang, W., Dasmahapatra, K., Mallet, J., Moreira, G. & Kronforst, M. (2016). Genome-wide introgression among distantly related *Heliconius* butterfly species. *Genome Biol*, **17(25),** doi 10.1186/s13059-016-0889-0.

APPENDIX I. SUPPLEMENTARY TABLES AND FIGURES

Table S1. Pairwise F_{ST} values among populations of *O. sylvatica.*

	Lita	Palo Amarillo	Tundaloma	Alto Tambo	Bilsa	Pto. Quito	Mediania	Quingue
Lita	0							
Palo Amarillo	0.02066	0						
Tundaloma	0.02876	0.00921	0					
Alto Tambo	0.03399	0.01692	0.01861	0				
Bilsa	0.02913	0.04358	0.05314	0.05382	0			
Pto. Quito	0.07252	0.05324	0.04654	0.04378	0.066	0		
Mediania	0.04171	0.04957	0.04066	0.04814	0.053	0.029	0	
Quingue	0.08234	0.09407	0.08676	0.09514	0.051	0.10597	0.07353	0

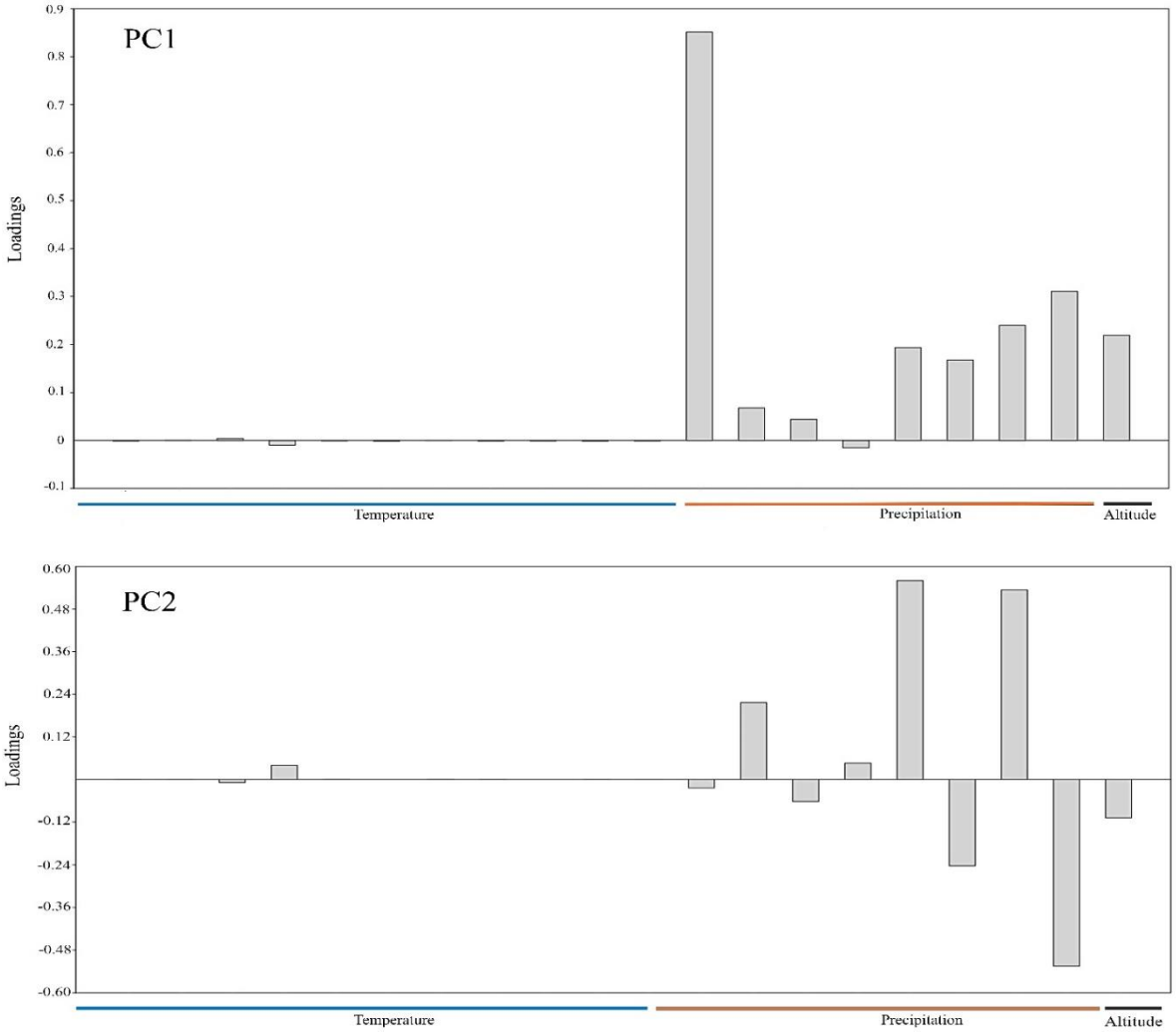

Figure S1: Loading of the two first components of the PCA analyses of the environmental variables. Precipitation is the factor that explain most the variation, followed by the altitude.

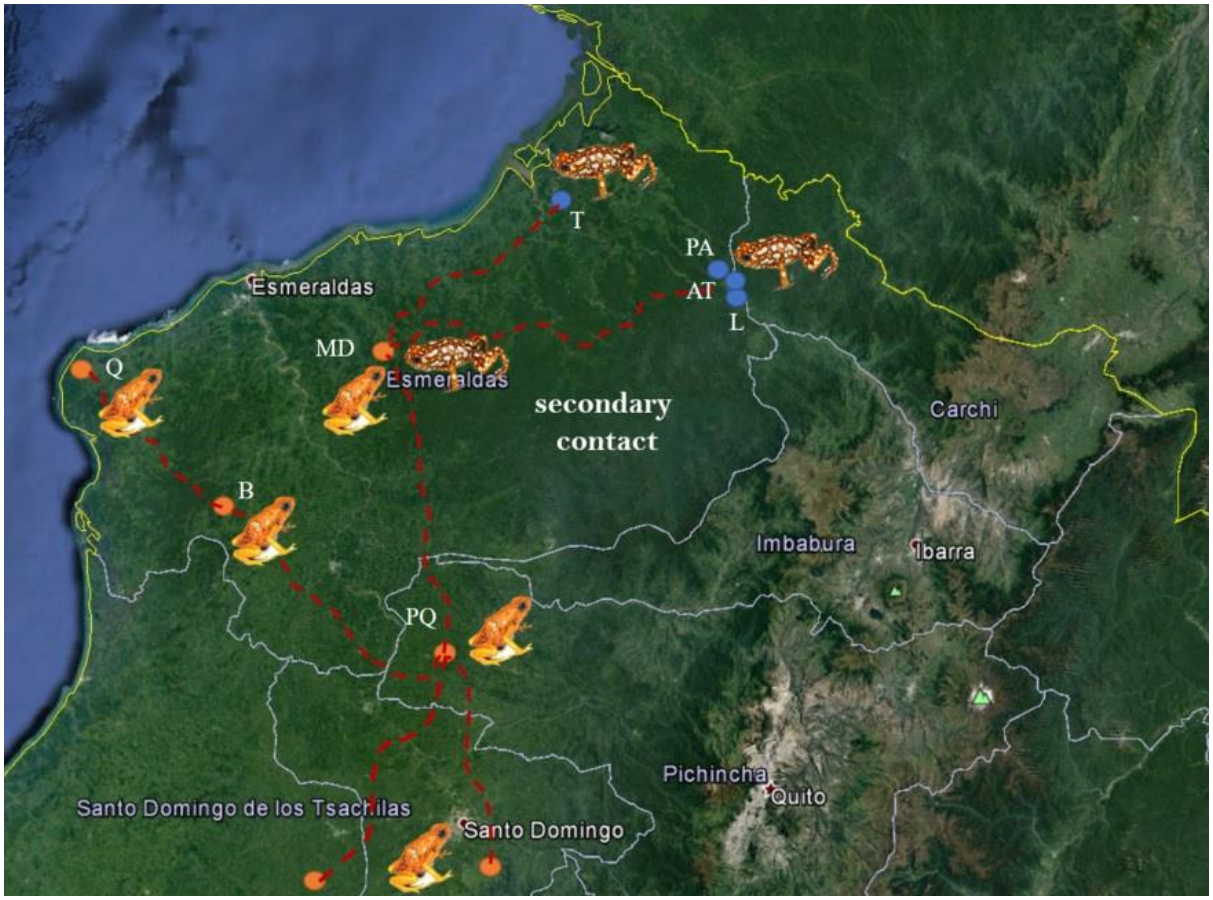

Figure S2: Map shows the presumable scenario of secondary contact between North and South populations at the beginning of present interglacial period.